THE UNIVERSITY
of EDINBURGH

Epistemic Inequality Reconsidered: An Inquiry into Epistemic Authority

MICHEL CROCE

Doctor of Philosophy
The University of Edinburgh
2020

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where states otherwise by reference or acknowledgment, the work presented is entirely my own.

_____ _____

Date Michel Croce

Abstract

Epistemic inequality is something we face in our everyday experience whenever we acknowledge our epistemic *inferiority* towards some and our epistemic *superiority* towards others. The *negative* side of this epistemic phenomenon has received due attention in the context of the debate on epistemic injustice: whenever an epistemic subject deflates the credibility of another or fails to recognize their authority qua knowers, unjust(ified) epistemic inequality is easily produced. However, this kind of inequality has an important (yet until now recently underestimated) *positive* side, as it can amount to a fundamental opportunity for the subject to achieve epistemic goods—such as knowledge, understanding, and intellectual virtues—from those who are epistemically superior to them.

This thesis inquires into the key features of the interaction between non-peer agents to provide a compelling account of epistemic inequality as a *positive* phenomenon, that is, as an inescapable and reliable source of epistemic goods. To achieve this overarching aim, the proposed analysis shall address three main questions, one conceptual, one normative, and one practical. The *conceptual* question asks what it takes for an epistemic agent to be epistemically superior to someone else. The *normative* question investigates the rational boundaries of the interaction between non-peer agents. From the standpoint of the epistemically superior agent, I examine the permissibility of epistemic paternalism, namely whether it can be legitimate for one to interfere with another's agency for the epistemic benefit of the subject interfered with without consulting them on the issue. From the standpoint of the epistemically inferior agent, I examine which attitudes it is rational to have towards individuals (whom we have reason to consider) epistemically superior to us. The *practical* question is about the implications of this analysis for a paradigmatic instance of interaction between non-peer subjects, such as the educational relationship between teachers and students.

As a response to the conceptual question, I shall offer a pluralistic virtue-based framework that allows individuating various ways in which an individual or a collective agent—such as groups, committees, and research teams—can be epistemically superior to another. In particular, I shall distinguish between experts, authorities of belief, and authorities of understanding.

As a response to the normative question, I argue that epistemic paternalism is permissible only insofar as the interferers are epistemically better positioned than the subjects interfered with, and I outline an account of what I call a "virtuous paternalist interferer". Furthermore, I argue that there are cases in which it is rational for an epistemically inferior subject to set aside whatever belief they might have and defer to the epistemically superior interlocutor—thereby endorsing a limited version of the so-called *preemptionist view* of epistemic authority.

Finally, as a response to the practical question, I shall propose an exemplar-based account of education, explain how it can be employed to educate the young to intellectual virtues, and defend the view from several objections raised against virtue-based approaches to education.

Acknowledgments

The opportunity to work as part of the Edinburgh wing of the *Diaphora* Innovative Training Network (under the European Union Research and Innovation Programme *Horizon 2020*, grant agreement n. 675415) made it possible to complete this thesis and provided financial support for my research throughout this journey.

I am hugely indebted to a considerable number of people, whose help and support have been fundamental to the completion of this work. The first mention goes to my supervisors, Duncan Pritchard and Aidan McGlynn, who managed to work with a student who has (hopefully) worked enough, but has surely taken too long to put together the various pieces of his work. Duncan Pritchard rescued me right when I was about to drown as a visiting graduate student from an Italian institution with nice dreams and no publications. I shall thank him forever for giving me a "second" chance. Among so many other things for which I am grateful to him, I will only mention the amount of written work from me he had to read and the suggestions as to how to get published and build an academic career. The outcome of this journey is my responsibility. But the premises of this journey are largely his doing. I am thankful to both Duncan Pritchard and Aidan McGlynn for patiently and relentlessly encouraging me to work without raising any doubts or worrying when I would fall silent for too long or was in the midst of writing papers and forgetting the dissertation project. Their supervision strategy made it possible to complete this journey while considerably improving my CV. Aidan McGlynn's input helped me to think more carefully and clearly about the implications of my work for other areas of epistemological interest (and beyond).

A special mention goes to four other fantastic scholars and great professors, whose help has been decisive in the completion of this thesis. Jennifer Lackey encouraged me to keep working on epistemic authority and supported my research visit to the University of Edinburgh, which changed my philosophical path for the better. Sven Rosenkranz supervised my research during my visit to Logos (University of Barcelona) during a crucial time in which I was trying to get several papers published. His help thus made several achievements possible and his supervision and attitude towards philosophical inquiry have been exemplary in many ways. His discreet but vivid direction at Logos should be an inspiration for many philosophical environments. I had the good fortune to meet Alessandra Tanesini at a considerable number of philosophical events and receive invaluable help and input from her throughout this entire journey. I have hugely benefited from her intellectual generosity and epistemic care, and I hope that our paths will continue to cross in the future. Linda Zagzebski has always generously answered all of the questions I have had about her work and has been very kind in giving her time to me on multiple occasions around the world, not to mention the opportunity to visit her at Oklahoma University back in 2014.

My philosophical wandering in the past few years allowed me to receive important input from many scholars and colleagues. Among them, special thanks go to Mark Alfano, Guy Axtell, Jason Baehr, Maria Baghramian, Heather Battaly, Amiel Bernal, Paul Blaschko, Kenneth Boyd, Fernando Broncano-Berrocal, Angelo Campodonico, Adam Carter, Annalisa Coliva, Jan Constantin, Ana-Maria Cretu, Mario De Caro, Jeroen de Ridder, Patrick Denning, Adriano Fabris, Domingos Faria, Luca Fonnesu, Sanford Goldberg, Modesto Gomez-Alonso, Emma Gordon, Peter Graham, Thomas Grundmann, Rie Iizuka, Christoph Jäger, Klemens Kappel, Jason Kawall, Chris Kelp, Arnon Keren, Ian James Kidd, Kristján Kristjánsson, Lex Lai, Yuanyuan Liu, Federica Isabella Malfatti, Carlo Martini, Elinor Mason, Michela Massimi, Luca Moretti, Benjamin McMyler, Hadeel Naeem, Mara Neijzen, Anna Ortin, Michele Palmira, Tommaso Piazza, Nina Poth, Christian Quast, Nicholas Rebol, Shane Ryan, Ricardo Santos, Harvey Siegel, Mona Simion, Nancy Snow, Lukas Schwengerer, Natasza Szutta, Guido Tana,

Luca Tuninetti, Maria Silvia Vaccarezza, Jesus Vega-Encabo, Lani Watson, Silvan Wittwer, René van Woudenberg, Sarah Wright, and Joost Ziff.

Another special mention goes to two colleagues at the University of Edinburgh, Giada Fratantonio and Matt Jope, who had the misfortune of sitting next to me for quite some time and being witness to both my happy and more challenging days. Perhaps her Italian roots might have made it easier for Giada to cope with me during that period, but Matt has certainly been reckless in choosing not to run away from me during our common visiting periods at Logos (University of Barcelona). Somewhat independently of how they survived those experiences, my life has been blessed by their presence.

Portions of this thesis have been presented at a number of venues. Among others, I shall mention: the 2016 annual conference of the British Philosophy Postgraduate Academy at the University of Reading; a workshop on *Expertise and Expert Knowledge* at the University College Dublin in 2017; the sixth graduate conference of the Wiener Forum für Analytische Philosophie at the University of Vienna in 2017; a conference on *New Trends in Epistemology* at the University of Pavia in 2017; the third *Diaphora Workshop* on *A-priori Knowledge* at the University of Stirling in 2017; a *Summer School in Social Epistemology* at the Autonomous University of Madrid in 2017; the second annual conference of *Aretai – Center on Virtues* in 2017; the *Virtue and Moral Education* conference at the University of Gdansk in 2018; the fourth *Diaphora Workshop* on *The Structure of Warrant* at the University of Edinburgh in 2018; the *Fake Knowledge Conference* at the University of Cologne in 2018; a conference on *Contemporary Issues across Ethics and Epistemology* at the University of Pavia in 2018; the 2018 meeting of the *European Epistemology Network* held in Amsterdam; a workshop on *Epistemic Tools and Goods in Education* at the University of Pavia in 2018; an invited talk at the *LanCog Group* at the University of Lisbon in 2018; the *Good Judgment, Critical Thinking, and Intellectual Virtues in Academic Education Conference* at the Free University in Amsterdam in 2019; the 2019 edition of the *Bled Epistemology Conference*; and a workshop on *The Role of the Exemplar in Arendt and Spinoza: Insight for Moral Exemplarism and Moral Education* at the University of Malmö in 2019. I am grateful to the organizers of these events for having me there and to the audiences for their comments and suggestions.

Some of the material included in this thesis can be found in various forms in earlier articles. In particular, Chapter 2 is based on "On What It Takes To Be An Expert", *The Philosophical Quarterly* (2019); Chapter 3 is based on "Expert-oriented Abilities vs. Novice-oriented Abilities: An Alternative Account of Epistemic Authority", *Episteme* (2018); Chapter 4 relies on "For a Service Conception of Epistemic Authority: A Collective Approach", *Social Epistemology* (2019); Chapter 5 is based on "Epistemic Paternalism and the Service Conception of Epistemic Authority", *Metaphilosophy* (2018); and Chapter 6 relies on "Epistemic Paternalism, Personal Sovereignty, and One's Own Good", in G. Axtell and A. Bernal (eds.), *Epistemic Paternalism* (Rowman & Littlefield, 2020). Finally, Chapter 7 brings together material from: "Educating through Exemplars: Alternative Paths to Virtue", *Theory and Research in Education* (2017, co-authored with M. S. Vaccarezza); "Il potenziale educativo degli esemplari intellettuali", *Ethics and Politics* (2018); "Exemplarism in Moral Education: Problems with Applicability and Indoctrination", *Journal of Moral Education* (2019); and "Moral Exemplars in Education: A Liberal Account", *Ethics and Education* (2020a).

The final thought goes to my parents, Antonella and Giorgio, my brother Mattia, and my partner Letizia, whose support, care, and love gave me the energy every day to complete this journey and, more importantly, to be happy throughout it. No words will suffice to express my profound gratitude to them.

Contents

Chapter 1

Introduction

Few topics in the epistemological domain are as in need of a timely discussion as the one about expertise. The main reason why this is the case is that the notion of an expert has been receiving repeated attacks in a variety of domains. To mention just some sources of skepticism towards experts, climate scientists are under attack from climate change denialists, virologists are under attack from anti-vax movements,[1] the National Institute of Standards and Technology—which conducted an official investigation on the causes of the fall of the Twin Towers—has been under attack from conspiracy theorists of all stripes. The list could go on and on, but these examples should suffice to signify the widespread distrust in expertise that currently affects our societies.

Undoubtedly, the Internet and, more specifically, the democratization of knowledge has played an important role in fostering such skepticism towards the role of experts and the criteria for attributing expertise (Lynch 2016). Social media, in particular, has become one of the leading sources of information as well as misinformation[2]. On the one hand, they make pieces of allegedly relevant news appear in our news feed and flow without interruption. But, on the other, they trigger our proneness to cognitive biases and facilitate the rising of socio-epistemic structures that impede the distribution of some range of information (i.e., epistemic bubbles and echo chambers; Nguyen 2018). Furthermore, they increase our tendency to take our beliefs to extremes whenever we take part in a group of individuals who share the same view (i.e., group polarization; Sunstein 2017). This has an obvious impact on the relationship between laypeople and experts: cognitive biases, such as confirmation bias and desirability bias, make us prone to disregarding expert opinions when they contradict our own views on a given matter, while epistemic bubbles and echo chambers work in such a way that unwelcome expert views do not get heard and shared within a community.

However, as Tom Nichols has recently pointed out (2017), blaming the Internet or social media for what he calls *the death of expertise* is a mere shortcut that fails to get at the heart of

[1] At least, they were until the Covid-19 outbreak that started in early 2020. I shall return to this issue at the end of this thesis, in chapter 8.

[2] For a discussion on whether misinformation and fake news amount to information see, e.g., Croce and Piazza (forthcoming), Gelfert (2018), Pritchard (forthcoming).

the problem. A more promising way to get a grip on this phenomenon includes asking ourselves why the education systems of several countries care so much about ensuring that the students obtain high scores on tests but care much less about making sure that they acquire an understanding of the subject matter, not to mention critical thinking abilities.[3]

There is more to the reasons for the skepticism towards expertise. Another relevant aspect concerns the various ways in which the communication between experts and the general public can, and in fact does, fail. Two obvious reasons why this can happen refer to the experts' inability and unwillingness to make their competence accessible to laypeople. As regards the concern with the experts' abilities, in chapters 2 and 3 I shall defend the thesis that being able to give a successful contribution to the acquisition of knowledge and being able to distribute knowledge within a community are very different things, and there is no reason why we should infer from the fact that someone possesses the former that they also have the latter, and vice versa. In short, this is why our societies not only need good scientists and scholars, but also competent disseminators and teachers. As regards the concern with the experts' willingness to engage in knowledge distribution, it seems evident that we cannot afford to allow experts to work in ivory towers or limit themselves to engaging with their restricted number of peers. Nonetheless, the distinction between experts and disseminators is also key to explaining why we do not want the former to waste their time in the attempt to serve a function, namely divulging information to the general public, that they might not be able to fulfill successfully.

Despite the current problems affecting the class of experts and their relationship with laypeople, the complexity of our social, financial, and political lives is such that we are still much in need of them and we ought to do all we can to ensure that our criteria to assess and confer expertise survive the skeptical threats and contribute to infusing due trust in those who deserve it. This dissertation has the ambition to offer a contribution in this direction: in particular, it aims at showing what epistemology can do to improve our understanding of expertise by advancing the discussion on its conceptual, normative, and practical dimensions at once.

As the title suggests, I will approach the question about expertise from a more general perspective, one that takes issue with the phenomenon of *epistemic inequality*. Each of us faces epistemic inequality in our everyday experience, whenever we acknowledge our epistemic *inferiority* towards some and our epistemic *superiority* towards others. As an epistemological topic, epistemic inequality crosses several contemporary issues in social epistemology. The *negative* side of this notion has generally received due attention in the philosophical debate and

[3] I will discuss the critical thinking approach to education in §7.4.2.

has typically taken two main forms: one has to do with whether human beings should be considered as having equal—better, comparable—sets of standard epistemic equipment; the other pertains to how our epistemic agency can generate unjust disproportions in the assessment of someone's epistemic capacities.

The question about our standard epistemic equipment connects modern epistemology with contemporary feminist standpoint theory: since Descartes and Spinoza, among others, it has been asked whether individuals have equal intellectual capabilities and are equally disposed to perform cognitive activities (e.g., Meeker 2013; Rorty 1996). Most modern philosophers seemed to assume *epistemic egalitarianism*, namely the thesis according to which the differences in epistemic capacities that human beings exhibit are far less than what they have in common, thereby justifying the conclusion that we are roughly equal as regards our standard epistemic equipment. This widely shared assumption has led contemporary epistemologists to endorse what Louise Antony famously called the *interchangeability thesis*, according to which any epistemic subject can be used to represent the rest (1995: 63).

The main genuine reason underlying this thesis is that it would be advantageous for our epistemological theories, arguments, and claims if we could rely on the idea that what applies to you as an epistemic subject can easily be adapted to my case as well as to the situation of any other epistemic agent. Nonetheless, this thesis has also led epistemologists to keep working with an idealized notion of an epistemic subject—call it *S*—whose social identity, personal interests, emotions, habits, virtues, and vices play no role in their knowledge acquisition processes. Feminist standpoint theory—but also virtue and social epistemology—challenged this ideal of a neutral and idealized knower, based on the assumption that knowledge is always situated and takes place within epistemic agents who, in fact, have social identities, preconceptions and biases, emotions, values, and interests. In particular, feminist standpoint theory became challenging to epistemic orthodoxy by showing that changing the subject, even while keeping every other epistemic factor fixed (e.g., the available evidence), might have a bearing on epistemic standing. In other words, feminist epistemologists like Antony (2018), Code (1991), Collins (1990), Harding (1991), and Longino (1990), among others, revealed that the egalitarian assumption rooted in the interchangeability thesis was, in fact, hiding a situated dominant assumption that the typical knower *S* was at least masculine, if not anything more specific and stereotypical than that.[4]

As I have anticipated, the other main problematic feature of the notion of epistemic inequality has to do with our actual epistemically relevant practices and, in particular, with

[4] See Grasswick (2018) for an overview of feminist social epistemology.

those practices that generate or foster what Miranda Fricker has famously called *epistemic injustice* (2007). In general, epistemic injustice is harm done to a person in their capacity as a knower or an epistemic agent (see, e.g., McGlynn 2019). This harm is typically produced by a (racial, gender, financial, cultural, etc.) prejudice and can assume two main forms. We are in the presence of *testimonial injustice* whenever "a speaker receives a deficit of credibility owing to the operation of prejudice in the hearer's judgment" (Fricker 2013: 1319) and in the presence of *hermeneutical injustice* whenever (a subject who belongs to) "a group which does not have access to equal participation in the generation of social meanings" suffers a disadvantage when "it comes to making sense of a significant area of their social experience" (*ibid*).

However, epistemic inequality has an important, yet underestimated, *positive* side, as the mere fact that a disparity in epistemic standing exists among epistemic subjects can constitute a fundamental opportunity for the epistemically inferior subject to achieve epistemic goods—such as knowledge, understanding, and intellectual virtues—from those who are epistemically superior to them (see Pritchard 2013a, 2015).

By exploring the attitudes which we are rationally permitted and/or required to adopt before experts or—as I shall define another fundamental category of epistemically superior individuals—epistemic authorities, this dissertation contributes to demolishing a second epistemological presupposition that we have inherited from the Moderns: the doctrine of *Knowledge Individualism* (Goldberg 2010). According to such doctrine, "all of the epistemic conditions on knowledge pertain to (states of mind and/or cognitive processes in) the knowing subject herself" (3). On this approach, acquiring knowledge is an individual business, an activity that involves single—and often isolated—epistemic agents relying on their sole cognitive faculties to achieving epistemic goods (Coady 2012: 27-28). Knowledge Individualism goes hand in hand with the idea that we should be wary of what other people believe and tell us, in that trusting one another would expose us to their potentially misleading or manipulative intentions. Furthermore, it explains why, since modern epistemology, perception, introspection, memory, and inferential reasoning have been considered reliable epistemic sources, while testimony has not. For these reasons, the project I will pursue here should be intended as a contribution to the field of social epistemology[5].

This dissertation takes issue with three main questions concerning the epistemology of expertise. The *conceptual* question asks what it takes for one to be epistemically superior to someone else. The *normative* question investigates the rational boundaries of the interaction

[5] For an overview of social epistemology's main questions and methodology, see, e.g., Goldman and Blanchard 2019.

between non-peer agents. The *practical* question inquires into the implications of this analysis for a paradigmatic instance of interaction between non-peer subjects, such as the educational relationship between teachers and students. The remainder of this introductory section is devoted to making a few terminological and conceptual clarifications for the sake of argument (§1.1) and further motivating the structure and the scope of this dissertation (§1.2).

1.1 Preliminary Remarks

Before getting started with the analysis of the selected issues surrounding the notion of cognitive expertise, it is helpful to clarify the use of some technical terminology that will accompany the reader throughout this entire thesis. The first clarification concerns the term *epistemic authority*. A good place to start is a famous article published by John Hardwig in the *Journal of Philosophy*, entitled "Epistemic Dependence" (1985): in this work, Hardwig challenges epistemologists to explain why any epistemic subject in standard psychological and physical conditions possesses a high number of beliefs for the truth of which they have no empirical evidence. His diagnosis was that epistemic subjects are justified in possessing beliefs that lack evidential support to the extent that those beliefs arise out of the epistemic dependence that epistemic subjects have—and acknowledge to have—toward more epistemically competent individuals, namely the experts.

Hardwig's thesis was far from trivial if we just bear in mind that not much had been done on the social dimensions of knowing at the time of him writing that article. Here is how he makes use of the notion of authority in the epistemic domain:

> If I am correct, appeals to *epistemic authority* are essentially ingredient in much of our knowledge. Appeals to the *authority of experts* often provide justification for claims to know, as well as grounding rational belief. At the same time, however, the epistemic superiority of the expert to the layman implies rational authority over the layman, undermining the intellectual autonomy of the individual and forcing a reexamination of our notion of rationality. The epistemic individualism implicit in many of our epistemologies is thus called into question, with important implications for how we understand knowledge and the knower, as well as for our conception of rationality (1985: 336; italics are mine).

Interestingly enough, this passage from Hardwig's paper shows that the expression *epistemic authority* was commonly considered as an attribute of experts or epistemically competent individuals. It also helps to illustrate the impact that any work on the topic of expertise has on the thesis of epistemic individualism I have discussed earlier.

Hardwig's terminological approach to the notions of expertise and epistemic authority has been adopted in further seminal works on this topic. Steve Fuller, for example, discusses the relationship between laypeople and experts as follows: "the layman is 'epistemically dependent' on the authority of experts for all but the beliefs on which he himself is an expert. […] Analytically, it may simply be a consequence of what it means to be an expert that one defers to its authority" (1988: 278). Similarly, in a passage discussing Philip Kitcher's work on epistemic authority, Alvin Goldman claims that "a scientist uses his own opinions about the subject-matter in question to evaluate a target scientist's degree of *authority*" (2001: 90; italics are mine).

Recently, Linda Zagzebski has proposed broadening the meaning of the term *epistemic authority*. Besides deploying this expression to refer to the normative power that the bearer of authority in the epistemic domain possesses (i.e., the expert), Zagzebski also refers to an epistemic authority as the epistemic agent—i.e., a person, a group, or an institution—that has authority in the epistemic domain. In this thesis, I will adopt Zagzebski's double usage of *epistemic authority* and, as will become clear in the chapters to follow, I take epistemic authorities and experts to identify two different categories of epistemically superior agents as regards the requirements and the function they are supposed to fulfill.

The second clarification concerns the term *expertise*. It should come as no surprise that the question about defining who an expert is occupies scholars from a wide range of disciplines, including epistemologists, philosophers of science, virtue theorists, sociologists, psychologists, and so on. This happens because we need to know whom to trust for acquiring knowledge in any given field, but also for figuring out how to behave in certain circumstances. In the latter case, having some sort of role model to look at seems to make things easier and reduce the probability of us getting things wrong by acting on our own.

One of the main concerns of those who discuss the credentials of expertise in psychology and the social sciences pertains to increasing "the chance that the process of coming to be called an expert will have more to do with the possession of real and substantive expertise" than with the identification of those who possess the reputation or the institutional markers of expertise (Collins and Evans 2007: 2).

Another important concern that has been highlighted in the works of brothers Stuart and Hubert Dreyfus (1991) on *practical* expertise has to do with individuating those elements that allow identifying who masters a skill to the point that they can be considered practical experts. Two peculiar features of this model are worth mentioning in the context of this dissertation. The first is that the process of skill acquisition moves from initial stages in which one acquires experience by following and applying simple rules to more proficient stages in which one develops an intuitive form of decision-making and navigates through a wide range of non-standard situations without losing control of their skill. The second is that "normally an expert does not *solve problems*. He does not *reason*. He does not even act deliberately. Rather he spontaneously does what has normally worked, and naturally, it normally works" (1991: 235). As a result, experts often struggle to account for what they know and can do, especially when it comes to explaining their abilities to novices.

To find a discussion of related topics in epistemology, we should look into the debate about knowing-how and, more specifically, about whether knowing-how boils down to knowing-that (see, e.g., Carter and Czarnecki 2016; Carter and Pritchard 2015a, 2015b; Ryle 1949; Stanley 2011; Stanley and Williamson 2001). It is surprising to notice that despite a clear—though partial—overlap of issues in the debates about practical expertise and know-how, we are still in need of comprehensive approaches to this intriguing and important philosophical topic. Examples of what the epistemological discussion considers experts in knowledge-how include, among others, chicken sexers (e.g., Masui and Hashimoto 1933; Lackey 2016; Zagzebski 1996), art experts who can identify fake pieces of art without being able to justify their verdict (Lackey 2016), and balsamic vinegar testers in Modena (Penco 2014). All these cases feature individuals who have a practical ability that allows them to reliably execute a performance, lack the capacity to provide an adequate account of what makes them able to perform their task successfully, and can transfer their know-how to novices through practical guidance—i.e., through a method based on attempts and feedback.

However, as Goldman makes clear, many if not most cases of practical expertise include—if not require to be grounded in—a superiority in propositional knowledge. To take just a couple of examples, Goldman mentions *repair* domains and *design* domains (1991: 129-130): both plumbers and engineers deploy their know-how to do their job but, in the vast majority of cases, they are able to provide true answers to questions concerning their methods, the problems they are fixing, or the procedures they are following to build something. This is because their know-how is grounded in a great deal of propositional knowledge about their domain and this

knowledge can—setting aside particular limitations of individual experts—be distributed to laypeople.

The Dreyfus and know-how models of expertise have attracted the attention of those virtue theorists and moral psychologists who take expertise in the moral domain to be a matter of mastering a skill and making (morally) good decisions rather than possessing moral knowledge and/or understanding (e.g., Annas 2011; Narvaez and Lapsley 2005; Stichter 2018). Although the model of expertise I shall defend includes some sort of a skill component and establishes a distinction between being an expert and being good at divulgating information, the focus of this dissertation is what Goldman calls *cognitive* expertise (2001: 91) rather than practical or performative expertise (see also Watson 2018: 40). In other words, I shall inquire into what it takes to be someone whom laypeople can trust if they want to improve their epistemic wellbeing, that is, if they want to acquire knowledge, understanding, or intellectual virtues.

The third and final clarification concerns the boundaries of my project. An important aspect of philosophical inquiries into the concepts of expertise and epistemic authority has to do with establishing reliable criteria through which laypeople can individuate people they can trust in a given domain. This dissertation does not purport to advance the discussion on epistemic authority in this respect. Rather, its main goal is to contribute to the debate about defining the epistemic features that make someone an expert in a given domain. Since this will be the main concern of the following chapters, let me briefly offer a few general considerations about the former issue.

From what Collins and Evans (2007), Goldman (2001), and Hawley (2012) have argued in their works on the topic, it is possible to extract at least four key markers of expertise that are somewhat available to laypeople:

(i) *Track record*: what an expert has successfully achieved in terms of making sound judgments in their field;

(ii) *Experience*: the familiarity an expert has with the questions that arise in their domain and the methodologies that are typically deployed to address such questions;[6]

[6] The relevance of *experience* as a marker of expertise has been recently highlighted by Matt Stichter in his discussion of *practical* expertise, when he contends that "[e]xperience not only changes how experts view a situation, it also enables them to efficiently and effectively respond to the situation" (2015: 110). This equates to what psychologists Darcia Narvaez and Dan Lapsley describe as "perceiving the affordances": "Because they have more and better organized knowledge in a domain, experts perceive things differently than do novices. They perceive different affordances" (2005: 150-151).

(iii) *Credentials*: the social status an expert has acquired allows laypeople to identify—reliably yet not infallibly—who knows what in a given domain. This marker relies on the idea that the acquisition of social credentials of expertise is grounded in reliable mechanisms and procedures involving several expert members of a community that assess and verify that a candidate possesses the required epistemic credentials on behalf of the entire community;

(iv) *Access to information*: the information an expert has often is not available to the majority of people, both because experts can make sense of data that "do not speak" to laypeople and because experts have a privileged epistemic position compared with laypeople (e.g., eyewitnesses in trials).

It is worth pointing out that this list of markers of expertise does not amount to a set of necessary and/or sufficient conditions for allowing laypeople to identify whom they should trust in a domain. Quite to the contrary, there seems to be some disagreement among the aforementioned scholars about the relevance of these indicators. All these scholars agree that the track record is a necessary requirement for identifying expertise, yet they have different takes as regards the other markers. According to Collins and Evans (2007: 67), credentials are an unreliable marker because they are unable to identify a number of kinds of experts, basically all the forms of connoisseurship that do not involve professional roles (e.g., expert violin players and individuals who are fluent in a language). Conversely, Hawley (2012: 68-71) argues in favor of credentials as a marker of expertise but does not mention experience, which appears to be a key marker on Collins and Evans' view.

Setting aside further considerations about the best way to cash out the markers of expertise, let me conclude by stressing that philosophical discussion on this topic is important especially because these markers should allow us to assess the expertise of our interlocutors before what Goldman calls *esoteric statements*, that is, in those circumstances in which the truth-values of the interlocutor's statements are inaccessible to us due to a lack of competence or possibility to assess their truth-values (2001: 94).

1.2 The Structure of the Thesis

As anticipated, this dissertation is built around three main questions concerning the notions of epistemic authority and cognitive expertise. This section offers a brief overview of the answers

I will put forth and motivate in this thesis and sets out the structure of the arguments of the following chapters.

THE CONCEPTUAL QUESTION

The conceptual question asks what it takes for one to be epistemically superior to someone else. I will defend a pluralistic theory according to which epistemic authority varies along two main axes: quantity and quality. The issue with quantity is that of figuring out whether expertise is a comparative matter—A is an expert for B iff A is somewhat epistemically superior to B—or a question of reaching an objective threshold, regardless of any inter-subjective dimension. The answer I will offer in chapter 2, in short, is that a compelling notion of an expert has to encompass both dimensions at once: below a certain threshold of epistemic qualities talking about epistemic authority is inappropriate but, at the same time, a comparative dimension is ineludible if we care about individuating those epistemic subjects whose competence can bring about epistemic benefits for the community.

As regards quality, the differences between various kinds of epistemically superior subjects pertain to both the *epistemic features* that make one better positioned than another and the kind of *service* that such individuals can provide for the benefit of the epistemic community. The aforementioned distinction between researchers and teachers incorporates both elements at once. Researchers constitute the paradigmatic example of what I shall call *experts*, namely those individuals who possess better knowledge than the majority of people in a given domain as well as a range of intellectual virtues—call them *research-oriented abilities*—which put them in a position to give a significant contribution to the epistemic progress of such a domain. Instead, teachers are the paradigmatic case of *epistemic authorities*, namely those individuals who are epistemically better positioned than their interlocutor(s) in a given domain and possess a range of intellectual virtues—call them *novice-oriented abilities*—which put them in a position to help the interlocutor(s) improve their epistemic standings in such a domain.

This difference in epistemic qualities is reflected in the kinds of functions these epistemic subjects are supposed to fulfill, as will become clear in chapters 2 and 3. On the one hand, we expect from experts that they help the community improve its knowledge and understanding of a given domain, by ensuring that reliable information is introduced in the community and unreliable information is left out—in fact, this amounts to performing a research-oriented or domain-oriented function, or in John Greco's words, a *gatekeeping* function of testimonial practices (2015: 283). On the other, we expect from epistemic authorities that they help the

epistemic welfare of a community by ensuring that as many members as possible acquire knowledge and understanding of the information available in a given domain—in fact, this amounts to performing a novice-oriented function or, in Greco's words, a *distributing* function of testimonial practices (284). As the reference to teachers should have made clear, it is key to an effective fulfillment of this function that epistemic authorities tailor their service to the specific possibilities of their interlocutor(s), depending on their background knowledge and epistemic capabilities.

Chapter 4 will explore the implications of the conceptual question for collective or group epistemology: my main aim in this chapter is to show that the proposed accounts (and the functions) of an expert and an epistemic authority can be applied to (and can be performed by) collective bodies such as research teams, companies, and institutions.

THE NORMATIVE QUESTION

The normative question that this dissertation explores concerns the rational boundaries of the interactions between non-peer agents. The focus will be on two specific issues, one pertaining to the norms that the novices should follow when they receive information from experts or epistemic authorities, and the other pertaining to the kinds of interventions that epistemically superior agents can undertake for the (epistemic) benefit of the novices from an epistemic point of view.

The former issue will be discussed in chapter 3: I shall defend what I will call a *limited-preemptionist view* of epistemic authority, according to which (only) in specific circumstances it can be rational for novices to let the judgments of the epistemically superior subjects preempt their own and therefore defer to experts and epistemic authorities. However, it will also become clear that, in the majority of cases, everyone should factor the opinion of epistemically superior individuals in a larger process of rational evaluation which takes into consideration also one's own evidence in favor of (or against) a given proposition.

The latter issue will be considered in chapters 5 and 6, which are entirely devoted to the topic of epistemic paternalism. Epistemic paternalism is the thesis that it can be legitimate for one to interfere with another's agency for the epistemic benefit of the subject interfered with without consulting them on the issue. In the recent discussion about this phenomenon, most efforts have been devoted to elucidating when and how paternalistic interferences are epistemically legitimate. My contribution to the debate is twofold: on the one hand, I will inquire into *who* is rationally entitled to undertake epistemically paternalistic interferences, and

in virtue of which features one inherits this entitlement (chapter 5); on the other, I will attempt to motivate epistemic paternalism by exploring *why* those interferences are justified (chapter 6).

More specifically, chapter 5 takes issue with Alvin Goldman's thesis that paternalistic interferences are justified insofar as the interfering subject is an *expert*. In the attempt to shed light on what are the required epistemic credentials of paternalist interferers, I shall argue that neither experts nor epistemic authorities are, as such, in an ideal position to interfere with laypeople's epistemic agency for their own welfare. Rather, I defend the view according to which not only are paternalist interferers supposed to be epistemically superior to the subject interfered with, but they also ought to deploy a wide range of novice-oriented abilities.

Chapter 6 addresses a challenge that has been recently raised by Emma Bullock (2018). In short, epistemic paternalists attempt to show that interfering with someone's epistemic agency for their epistemic benefit cannot be reduced to a mere instance of general paternalism, in that there is something specific about the epistemic nature and goal of the intervention. Yet they seem unable to handle the fact that any paternalistic interference—hence, including interventions in the epistemic domain—clashes with the ineludible value of one's personal autonomy, or so Bullock argues. In other words, epistemic paternalists cannot have their cake and eat it too. For epistemically paternalistic interferences to be—at least in principle—justified, their value has to lie in the fact that they contribute to one's overall—rather than epistemic—wellbeing. Thus, to address the problem with personal autonomy, epistemic paternalists seem committed to concede that there is nothing inherently "epistemic" about the interferences they are willing to consider legitimate, in that their value merely lies in the fact that they outweigh one's personal autonomy but contribute to one's overall wellbeing.

In chapter 6, I shall take stock of Bullock's dilemma and argue that epistemic paternalists still have one string to their bow: if they renounce interpreting the notion of personal autonomy as mere personal sovereignty and adopt alternative accounts of autonomy (e.g., the model of autonomy as a condition illustrated in Feinberg 1989), it becomes clear that some paternalistic interferences are justified because they constrain one's personal sovereignty in the short run to improve one's personal autonomy—and overall wellbeing—in the long run.

THE PRACTICAL QUESTION

The practical question concerning epistemic authority asks what can be done to improve the epistemic quality of the relationship between non-peer subjects in paradigmatic settings such

as education. The proposed analysis will shed light on an important—yet currently underestimated—epistemic good that epistemic authorities can help the novices acquire, namely intellectual virtues. As will become clear in chapter 7, I do not endorse the claim that epistemically superior individuals can pass along epistemic virtues to novices. Nonetheless, the inquiry I shall conduct in the field of character education will reveal that novices can acquire or improve their intellectual character traits by admiring and attempting to emulate the epistemic conduct of virtuous agents, whom I shall call *epistemic exemplars*. I will appeal to Zagzebski's exemplarist theory in the moral domain to offer what I take to be the most promising version of an exemplar-based account of education to intellectual virtues and I will defend this account from a couple of central objections that have been raised against this educational approach.

Standard exemplarist theories have it that an exemplar is a paradigmatically good agent, that is, an exceptionally virtuous one. Traditional accounts of exemplarity in the moral and the epistemic domain distinguish between at least two categories, namely heroes and saints. The former are taken to be exceptionally courageous or, at least, exceptionally virtuous in one specific respect, while the latter are supposed to display all the virtues (e.g., Blum 1988; Wolf 1982). The *liberal* account of epistemic exemplarity that I will offer departs from this tradition and distinguishes itself in at least two respects. More specifically, I shall argue that, if what we are concerned with is the educational applicability of exemplarist theories, such traditional requirements as the exemplar's virtuousness and imitability are unnecessary. Chapter 7 will explain that, for one thing, less-than-virtuous individuals—which I will call *enkratic exemplars*—can well serve the purposes of epistemic exemplarity for vicious and akratic novices. For another, virtuous yet not imitable individuals—which I will call *injustice illuminators*—deserve to be considered as epistemic exemplars because they help most members of an epistemic community become aware of some unjust relationship or social structure, although their admirable behavior might not be easily emulated by those who are not members of oppressed groups.

The analysis of these interesting cases will support the conclusion that for an exemplar-based approach of intellectual virtue formation to be educationally effective, the liberalizing turn I propose is only a first step of a long journey.

Chapter 2

On What It Takes to Be an Expert

Abstract. This chapter tackles the problem of defining what a cognitive expert is. Starting from a shared intuition that the definition of an expert depends upon the conceptual function of expertise, I shed light on two main approaches to the notion of an expert: according to *novice-oriented accounts* of expertise, experts need to provide laypeople with information they lack in some domain; whereas, according to *research-oriented accounts*, experts need to contribute to the epistemic progress of their discipline. In this chapter, I defend the thesis that cognitive experts should be identified by their ability to perform the latter function rather than the former, as novice-oriented accounts, unlike research-oriented ones, fail to comply with the rules of a functionalist approach to expertise.

2.1 Introduction

Recent works in social epistemology have shown how challenging it is to define what it takes for one to be a cognitive expert in some field. Yet it seems important to be able to distinguish those trustworthy subjects who contribute to the epistemic welfare of our communities from those who do not deserve our epistemic credit and resources.

Alvin Goldman has been trying to settle the issue for more than two decades (see 1991, 2001, 2018), but his view of expertise has received persistent criticism. Some of Goldman's critics (e.g., Coady 2006; Scholz 2009) accept his general strategy of reaching a satisfactory definition of an expert through conceptual analysis, but question some—or, every—condition of his account. Others reject the methodology of Goldman's project and suggest alternative strategies for analyzing the notion of an expert, such as a Wittgensteinian family-resemblance route (Majdik and Keith 2011), Edward Craig's practical explication approach (Quast 2018), and Nelson Goodman's symptom approach (Scholz 2018).

In this chapter, taking up insights from these approaches and from Goldman's most recent attempt to "advance a definition of expertise that … is in roughly the same territory of some of his critics" (2018: 3), I will demonstrate that a conceptual analysis still has an important contribution to make. Starting from the intuition that the definition of an expert depends upon

the epistemic function they fulfill in our communities,[7] the chapter sheds light on two different roles of an expert, i.e., that of improving laypeople's epistemic condition—broadly construed—and that of contributing to the epistemic progress of their discipline. The overall aim of this chapter is to show that cognitive experts should be identified by their ability to perform the latter function, rather than the former.

To accomplish this goal, in section 2.2 I introduce the two rival accounts of an expert, which I respectively name the *novice-oriented account* and the *research-oriented account*. Section 2.3 is devoted to elaborating on their conditions in order to make sure we are working with the most convincing version of each account before providing an argument against the novice-oriented approach, which will be offered in section 2.4. In section 2.5, I shall motivate the superiority of the research-oriented account over the rival view. Finally, section 2.6 explores some consequences of the argument for the social function of experts.

As should be evident, this chapter contributes to answering the Conceptual Question that lies at the grounds of this thesis. A complete answer to this question will require a more in-depth analysis of what it takes to be an epistemic authority, which I shall offer in the following chapter, as well as an inquiry into the epistemic authority of collective agents, which will be put forth in chapter 5.

2.2 Two Accounts of Expertise

According to Goldman's (2018) and Quast's (2018) recent diagnoses, the epistemological debate on the notion of the cognitive expert is based on a fundamental methodological assumption, according to which a compelling definition of an expert has to reflect their *function* within an epistemic community. More specifically, Goldman contends that such definition should highlight "what expertise *is* by reference to what experts can *do* for laypersons by means of their special knowledge or skill" (2018: 3). For these reasons, a functionalist approach to the notion of expertise should provide (i) a *functional definition* of an expert, and (ii) a *possession condition* that individuates what it takes for an epistemic subject to fulfill that function (4).

As regards (i), Coady (2012) and Goldman (2018) accept what I call a *novice-oriented* approach. According to them, "the main function of expertise is sharing some knowledge for the benefit of someone else", where the recipient is often intended as a layperson or a novice

[7] See Goldman (2018) and Quast (2018).

(Quast 2018: 13). With Coady's words, experts are supposed to fulfill "the practical role of the expert", that is "being someone laypeople can go to in order to receive accurate answers to their questions" (2012: 30). The role of experts suggested by Coady is reflected in Goldman's functional definition, according to which a subject S is an expert in domain D if and only if:

> [NO-CAP] S has the capacity to help others (especially laypersons) solve a variety of problems in D or execute an assortment of tasks in D which the latter would not be able to solve or execute on their own. S can provide such help by imparting to the layperson (or other client) his/her distinctive knowledge or skills. (2018: 4)[8]

As regards (ii), that is the cognitive states that allow one to fulfill this functional definition, Goldman recommends a veritistic or "truth-linked" possession condition, according to which an expert is someone who possesses more accurate information than most people do in a domain. Specifically, on his view a subject S is an expert in domain D if and only if:

> [T-LC1] S has more true beliefs and fewer false beliefs in propositions pertaining to D than do most people. (4)[9]

In sum, the novice-oriented account defines expertise as the capacity specified by NO-CAP and suggests T-LC1 as a necessary and sufficient characterization of what allows one to possess that capacity.

Interestingly enough, some earlier remarks by Goldman allow us to suggest a different account of an expert. Even though when he wrote his 2001 paper Goldman did not have to face objections to his approach based on conceptual analysis, he already had in mind that "expertise features a propensity element as well as an element of actual attainment" (92). The propensity element in his less recent view amounts to "an ability to generate new knowledge in answer to questions within the domain" (91) while the element of actual attainment amounts to an expert's possession of "more true beliefs and/or fewer false beliefs than most people do in D" (91). The distinction between propensity and actual attainment can now be adapted to his recent functionalist framework and offer grounds for what I call a *research-oriented account.*

[8] I replace Goldman's original label "CAP" with "NO-CAP" to stress that on this account experts are required to fulfill a *novice-oriented* function, and make it easy to contrast it with the *research-oriented* function I shall introduce in §2.5.

[9] I also replace Goldman's original label "TL1" with "T-LC1". The expression "belief in" is usually understood as involving an element of trust. However, that is not the connotation Goldman gives to this condition for expertise. I stick to his own connotation for the sake of consistency with his wording.

Specifically, the propensity—that is, disposition-based—element provides us with a research-oriented functional definition of expertise, according to which a subject S is an expert in domain D if and only if:

> [RO-CAP] S has the capacity to contribute to the epistemic progress of D. S can provide such help by offering true answers to the questions under dispute in D.

In contrast, the element of actual attainment gives us a slightly different version of Goldman's recent truth-linked possession condition, which I shall call T-LC2, according to which a subject S is an expert in domain D if and only if:

> [T-LC2] S possesses more true beliefs and/or fewer false beliefs than most people do in D.

Thus, the research-oriented account defines expertise as the capacity specified by RO-CAP and proposes T-LC2 as a necessary and sufficient characterization of what it takes for one to possess this capacity.

A quick comparison of the two accounts suggests a couple of remarks. First, notice that T-LC2 not only grants that someone can be an expert by having more true beliefs *and* fewer false beliefs than most people do in D, but also, and more importantly, that one can be an expert by having more true beliefs *or* fewer false beliefs than most people do in D. I take T-LC2 to be more problematic than T-LC1, but I shall discuss this in §2.3.1.

Second, someone might take the fact that two very similar versions of the truth-linked requirement serve as possession conditions on allegedly alternative accounts to be a sign that Goldman in fact thinks an expert to be someone who fulfills both a research-oriented and a novice-oriented function. Specifically, one might argue that if the propositions pertaining to a domain of expertise D include not only propositions at the object level of D but also propositions about the processes, methods, and skills the use of which is likely to yield true beliefs in D, then satisfying the veritistic condition on the novice-oriented account will typically result in the capacity to generate new knowledge in D (i.e., RO-CAP), at least in non-empirical domains.

I do not buy into this argumentative line, for at least two reasons. First, I think it is unsuccessful: as I will show in §2.5, there are ways in which one can have more true beliefs than most people do at all these levels of D and yet lack the capacity to contribute to the progress

of D. Second, no matter what the most faithful interpretation of Goldman's thought over the years is, we should still keep the two accounts separate for the sake of argument. If my arguments hit their target, we can conclude that the research-oriented view can account for what it takes to be an expert in Goldman's recent functionalist approach while the novice-oriented view fails to do so. Therefore, let us work on the assumption that neither view should be reduced to the other and examine the two accounts in detail.

2.3 Toward a Compelling Novice-Oriented Account of an Expert

The analysis of the novice-oriented account of an expert concerns the plausibility of the truth-linked condition and its relationship with the functional definition of expertise expressed by NO-CAP. Before starting to discuss the plausibility of the possession condition, it would be wise to address a preliminary issue that might affect our understanding of the notion of an expert. Some might be inclined to think that the truth-linked requirement makes Goldman's account *comparative*, as the attribution of expertise to someone partially depends on the reliability of most members of their epistemic community. Goldman does not deny that the notion of an expert has a comparative dimension that we refer to when we say that, for example, Anna is more of an expert in contemporary history than Nick, or that Giada is more of an expert in psychology than anyone else in the room. However, he adds a *non-comparative clause* (2001: 91), or *threshold condition* (2018: 5), to this account, according to which "[b]eing an expert is not simply a matter of veritistic superiority. Some non-comparative threshold of veritistic attainment must be reached" (2001: 91), even though this threshold might be difficult to determine.[10] Thus, on Goldman's view the notion of an expert requires a minimum threshold of true beliefs in D below which someone should not be considered expert, no matter if they are better informed on D than most people in their (or other) communities.

Coady rejects the threshold condition, considering it unnecessary for a novice-oriented account. The debate about the necessity of this requirement will be discussed in §2.3.2. Before

[10] In recent work (2018), Goldman has developed this thought and maintains that the non-comparative clause could be added to the definition of an expert as follows:
[TL-C*] S is an expert about domain D if and only if:
(A) S has more true beliefs (or high credences) in propositions concerning D than most people do, and fewer false beliefs; and
(B) The absolute number of true beliefs S has about propositions in D is very substantial.
He also admits that "the absolute condition stated in (B) is unclear" but contends that "whatever might be said to tighten this condition needn't affect the general type of approach proposed here" (5). Put it simply, it should remain a truth-linked, or veritistic, approach.

that, we shall discuss a specific objection against the possession condition, according to which T-LC1 is unnecessarily demanding, and offer a refined version of the condition that avoids this criticism (§3.1). Finally, we will challenge Goldman's veritistic project by arguing that an understanding-linked requirement better complies with NO-CAP than a truth-linked condition. A counterexample to standard novice-oriented accounts will be offered to support the claim that experts need to possess better understanding[11] than most people in their domain of expertise (§2.3.3).

2.3.1 The Expertise of Epistemically Negligent Subjects

As I pointed out in §2.2, the truth-linked requirement introduced in the research-oriented account (T-LC2) differs from the one displayed in the novice-oriented account (T-LC1). However, given that they share a similar structure and the same components, I will discuss both here with the goal of reaching a more robust version of the possession condition that can be employed in both accounts. More specifically, I shall argue that T-LC2 is unduly weak, whereas T-LC1 is unduly demanding.

On Goldman's view, in order for a subject to satisfy T-LC2 it is sufficient that she fulfills either of the following requirements:

[T-LCa] S possesses more true beliefs than most people do in D;
[T-LCb] S possesses fewer false beliefs than most people do in D.

Scholz (2009) and Coady (2012) argue that Goldman is wrong, because we are not willing to consider someone an expert if their superiority to laypeople in terms of reliability merely amounts to lacking some of the false beliefs laypeople have in a given domain, as required by T-LCb. A quick comparison of experts with epistemically negligent subjects should make clear that they are right.

In general, an expert has a much greater interest in the status and the epistemic progress of a given discipline than a layperson does. This interest inevitably comes with a greater risk of believing falsehoods in such a domain than laypeople do (see Coady 2012: 29; Scholz 2009: 193). Indeed, it may be the case that experts have both more true beliefs and more false beliefs about their subject matter than laypeople. Thus, experts might satisfy T-LCa without being in

[11] See §2.3.3 for a definition of "better understanding".

the position to satisfy T-LCb. In contrast, epistemically negligent people lack any interest in the epistemic progress of a domain and do not question their beliefs so frequently. Hence, they avoid any risk of acquiring false beliefs at the huge cost of not acquiring true beliefs. If we accepted T-LC2 then, epistemically negligent subjects would have a good chance of becoming experts simply by fulfilling T-LCb, particularly in those disciplines where there is widespread disagreement about the answers to the main questions of each domain.

This argument not only applies to epistemically negligent subjects, but also to those who doubt that we can acquire knowledge in some specific area of inquiry, i.e., local skeptics. These subjects can easily satisfy T-LCb: given that they doubt the possibility of acquiring knowledge in some domain, they avoid forming beliefs by suspending judgment in that domain. Thus, it is likely that they will possess fewer false beliefs than any other individual who assumes that knowledge can be obtained in that domain. According to Goldman's loose requirement, local skeptics will then be in the position to satisfy the possession condition on a research-oriented account of expertise by fulfilling T-LCb, as epistemically negligent people do.

But again, this result seems counterintuitive. If, for instance, we wanted to know more about the potential and the side effects of an experimental cure for cancer, it would be very odd to ask a convinced herbalist who treats ill people with herbal extracts and suspends judgment on everything related to traditional medical experiments. This subject will no doubt have fewer false beliefs on the effects of the experimental treatment than the oncologists who are testing the new medicine—in fact, she will have none since she withholds judgment. Nevertheless, it seems evident that we would rather consult the oncologists than the herbalist, as the former possess more true beliefs about the cure than the latter does.

From these considerations it follows that it should not be the case that T-LCb is a *sufficient* component of the truth-linked condition, in that even epistemically negligent lay subjects and local skeptics can satisfy it. Nor should it be the case that T-LCb is a *necessary* component of the truth-linked condition, because an expert who puts forth a new theory in D may entertain more false beliefs than laypeople in D.[12] This argument allows us to undermine both T-LC2 and T-LC1 at the same time. On the one hand, the fact that T-LCb is not a sufficient component of the possession condition proves that T-LC2 is too weak, because in general we shall not grant expertise to epistemically negligent laypeople. On the other hand, the fact that T-LCb is not a necessary component of the condition proves that T-LC1 is too demanding. As mentioned in §2.2, in order for a subject to satisfy T-LC1 it is necessary that she fulfill both T-LCa and T-

[12] Scholz contemplates an even more extreme scenario, in which a layperson might have more true and fewer false beliefs about D than an expert (see 2009: 193).

LCb. But, as I have just shown, the mere fact that an expert may possess more false beliefs in a domain than laypeople do should not undermine her expertise. Thus, there seems to be no reason for retaining T-LCb as a component of the truth-linked condition. Goldman (2018) seems to acknowledge the effectiveness of this objection in footnote, where he mentions the possibility of requiring that experts merely have "a comparatively high *ratio* of true to false beliefs to qualify as experts" (5, n.3). For our purposes though, it suffices to follow Coady and Scholz and strengthen the possession condition on either account of an expert by getting rid of T-LCb. The revised definition of an expert on the novice-oriented account then states that S is an expert in D if and only if:

[T-LCa] S possesses more true beliefs than most people do in D.

2.3.2 On the Necessity of a Non-Comparative Clause

The second criticism of the possession condition on Goldman's novice-oriented account pertains to the non-comparative clause introduced above. Coady contends that there is no need to add such a clause to the possession condition, because "if some individuals are significantly better informed than most people about a subject […] in their community, then […] they should be considered experts on it, whether or not they are well-informed from a God's-eye point of view" (2012: 29).

I agree with Coady that on a novice-oriented account there is no need of a "God's eye" perspective. If the function of expertise merely amounts to helping laypeople acquire new knowledge, then a comparative notion of an expert would easily fit the bill. To see why this is the case, let us consider a couple of examples featuring Goldman's idea of epistemic communities: that is, groups such as "a doctrinal community whose members devoutly and uncritically agree with the opinions of some single leader" (2001: 98), a "community of scientists" (103), and professional communities (105).

First, consider the case of Skyler, an excellent chemistry scholar who is presenting the results of her study at an International Chemistry Conference. Both Goldman and Coady can grant her the expertise she acquired through years of intense study and research. In fact, not only does she have the capacity to help laypeople acquire knowledge about chemistry, but she supposedly has more true beliefs than most people do about chemistry and she exceeds the threshold condition, wherever we stipulate that it has to be set.

Then, compare Skyler's case with that of Emma, a young member of a close-minded community led by a charismatic faith-healer who pretends to cure ill people with his spiritual powers. Suppose also that Emma challenges his opinion based on what she reads in an introductory textbook of medicine. In this case, Goldman's non-comparative clause would prevent Emma from being considered an expert in medicine, as she certainly fails to fulfill the threshold condition. But that clashes with the methodological assumption of his functionalist approach because Emma might well be in the position to help other members of her community acquire true beliefs in medicine. Coady's view avoids this problem: on his account, Emma can be an expert in medicine in her community because she is in the position to fulfill NO-CAP. As a matter of fact, Emma has acquired more true beliefs than the other members of the community by reading the introductory textbook and carefully observing the alleged treatments provided by the faith healer. Were someone to notice this fact, they would start to consider her an expert in the field and disregard the leader's advice.

It should now be evident that, unlike Goldman's view, Coady's comparative notion of an expert fits the methodological assumption introduced in §2.1 because it confers expertise to epistemic subjects based on the function that they can provide to other members of their communities. On the novice-oriented account, both Skyler and Emma can be considered experts because they have the ability to provide other members of their communities with information they lack, thereby fulfilling NO-CAP. Therefore, *pace* Goldman, there is no reason why we should constrain the novice-oriented account by imposing a threshold condition, when a comparative approach—such as Coady's—already ensures the accomplishment of the novice-oriented function.

2.3.3 Veritism and Understanding

Both objections considered so far concerned specific features of Goldman's truth-linked possession condition. However, his veritistic account of an expert incurs at least two further problems relative to its broad truth-linked scope.

The first problem arises from the fact that scientific progress is transitional by nature.[13] As an example, recall that before the Fifties most scientists believed that cigarettes did not contain harmful chemicals; now, we all know that smoking may cause cancer and other diseases. This consideration threatens Goldman's novice-oriented account because we are obviously unable

[13] This problem is raised by Watson (2018: 45), who calls it "the shoulders of giants problem".

to put ourselves in the "God's eye" perspective and evaluate who, among thousands of alleged experts, possesses a sufficiently large number of true beliefs in some domain. Goldman acknowledges this as "a common problem" (2018: 5), yet he holds that this is not enough to undercut the appropriateness of the veritistic approach. The mere fact that it's hard to find out who possesses true beliefs in some discipline does not entail that objective expertise cannot be found in that domain.[14]

Let us assume that Goldman's defense is successful. A second far more troubling issue puts pressure on his approach. Consider the following example:

> EXPERT BY HEART. Suppose that Denny, a young ski instructor with a special talent for remembering, decides to help his lazy brother Brett pass his physiology licensing examination. This difficult exam consists of a random selection of 100 multiple choice questions that each candidate receives on their laptop at the beginning of the exam session and has to answer in a very short amount of time. All candidates know is that the computer selects among 10000 questions that the examiners gave to each candidate together with the reading list. Denny offers to memorize the whole list of 10000 questions, with each respective correct answer, and to provide Brett with the list of correct replies ten minutes after receiving screenshots of the test. The plan works perfectly: following Brett's instructions, Denny selects the wrong answers to five questions on purpose, not to arouse the examiner's suspicion. With 95% of correct answers, Brett is the only candidate who passes the examination with distinction this year.

If we analyze Denny's profile under Goldman's requirements, we notice that Denny satisfies T-LCa by hypothesis, as he has acquired more true beliefs about physiology than most people do. Furthermore, it is reasonable to suppose that he goes beyond the threshold for being an expert in physiology: not only is he significantly better informed than most members of his community, but also the overall number of true beliefs in physiology he possesses is incredibly high. Finally, Denny fulfills NO-CAP by doing a great service to his brother, who did not prepare for the exam. Were other candidates to ask for his help in the future, he would be happy to do that, despite any considerations about the morality of his actions.

[14] Notice that the same problem affects any novice-oriented approach, including Coady's comparative view. For what is problematic in this respect is not the presence of a non-comparative clause, but rather the truth-linked condition on which novice-oriented accounts are based.

Nonetheless, it seems intuitively evident that Denny should not be considered an expert in physiology. There are at least two ways in which we can try to back this claim up. First, one might argue that Denny possesses lots of true, yet unjustified, beliefs: in fact, he has no reason to offer in support of these propositions about physiology. However, this argumentative line appears to be unpromising on closer inspection. On the one hand, it is unclear whether Goldman takes experts to possess not only a substantial number of true beliefs in D, but also a substantial number of *justified* true beliefs in D.[15] On the other, it can be argued that Denny acquires all this information about physiology from a reliable source, such as the list made by the board of examiners, and has no defeaters for it. Thus, many epistemologists would agree that Denny has acquired testimonial knowledge of these propositions.[16]

The second strategy to support the idea that Denny is not an expert appeals to the notion of *understanding* as a fundamental epistemic goal. This notion is open to different interpretations. Here I follow a well-established tradition that considers understanding as "grasping systematic connections among elements of a complex whole, or gaining insight into certain relations between items within a larger body of information" (Jäger 2016: 180).[17] Thus, I focus on the "objectual" dimension of understanding, as related to subject matters, theories, and bodies of information requiring explanation, rather than mere propositions.[18] I appeal to this notion to suggest that what should distinguish experts from laypeople is not merely the high number of true beliefs the former—unlike the latter—possess in some domain D. On the contrary, a more plausible measure of someone's expertise in D is their understanding of the background assumptions of D, the main questions arising within the discipline, and the available answers. In sum, experts are supposed to understand the relationships between the various components

[15] No mention of justification as a necessary requirement for expertise has been found in Goldman's 2001 definition, where he defines an expert as someone who has "extensive knowledge (in the weak sense of knowledge, i.e., true belief) of the state of the evidence" (92). In contrast, he explicitly admits the possibility of an *evidential (or justificational) approach* in contrast to, or in support of, a veritistic approach (see 2018: 6-7), but acknowledges that grounding the definition of an expert on a justification- or evidence-based requirement leads to several problems.

[16] Note that this problem is rooted in Goldman's particular form of veritism. Pritchard's (2019) virtue-centered veritism, for example, does not incur this problem in that, on his view, veritism is not understood in terms of maximization of true beliefs but rather in terms of virtuous inquiry, that is, seeking out the truth and being able to distinguish significant from insignificant truth. On this alternative veritistic view, Denny could not count as an expert precisely because he lacks the epistemic skills and intellectual virtues required to discriminate among the true bits he learns by heart.

[17] Further accounts of understanding can be found in, e.g., Grimm (2018), Kvanvig (2003), Malfatti (2019a), and Pritchard (2009).

[18] An important difference between objectual understanding and propositional understanding is that understanding a proposition *p*, e.g., understanding that my smartphone is broken, does not seem to differentiate much from knowing that *p*. Rather, understanding a subject matter x, or a domain D, e.g., understanding meteorology, cannot be reduced to merely knowing some propositions (see Kvanvig 2009). On the propositional model of understanding see also Grimm (2014: 330-ss). For a recent discussion of objectual understanding, see Carter and Gordon (2014).

of D. Going back to EXPERT BY HEART, it should be evident that Denny lacks expertise in physiology because he has no understanding whatsoever of the thousands of true beliefs he learned by heart. All he knows is the correct answer associated with each question of the examination list.

Therefore, I suggest that Goldman's truth-linked condition be replaced with the following understanding-linked condition. A subject S is an expert in domain D if and only if:

[U-LC] S possesses better understanding of D than most people do.

It is a widely shared view that understanding is matter of degrees. Following Elgin (2007), I take it to include the three following dimensions: *breadth*—S's ability to imbed some true belief pertaining to a domain into a more comprehensive set of beliefs concerning that domain; *depth*—the number of propositions and/or non-trivial inferential connections between propositions within S's network of beliefs; and *significance*—S's ability to appreciate the relevance of the beliefs included in D, not merely to recognize that they are true (36).

To give just an example of the dimensions of understanding, it seems safe to contend that Mary has a broader understanding of Real Madrid's football tactics than Felix if she is better able to imbed the beliefs they both possess about Real Madrid's offensive and defensive strategies into a wider set of beliefs concerning the tactics that coach Zinedine Zidane has adopted in the past as well as previous strategies that Real Madrid adopted with other coaches. Mary has a deeper understanding of Real Madrid's tactics than Felix if she grasps the mechanics of the team's defensive rotations and the give-and-go movements to build offensive strategies while he just displays a superficial understanding of *Los Blancos*' tactics. Finally, supposing that Mary and Felix are studying to become football coaches, we can argue that Mary has a more significant understanding of Real Madrid's tactics than Felix if she can better assess the importance of grasping the innovative strategies that Zidane introduced as the Real Madrid coach for her own epistemic background and professional development.

I consider the notion of "better understanding" to include all three dimensions just introduced, to the extent that if S is an expert in D, S has broader, deeper, and more significant grip on the relationship between true bodies of information constituting D than most people do.[19] Thus, adopting U-LC not only avoids the problem generated by EXPERT BY HEART; it also offers a more plausible rendering of the veritistic account Goldman aims to provide.

[19] Scholz (2018) introduces the idea of understanding as a "symptom of expertise" and admits a great variety of objects of understanding (34-36). He also concedes that understanding can vary along breadth and depth, but makes no reference to the dimension of significance, nor to Elgin's account. I am not sure what the relationship between

To briefly summarize, section 2.3 has been devoted to offering a more plausible version of the novice-oriented account of an expert originally proposed by Goldman. First, I stressed the advantages of amending Goldman's truth-linked condition as Coady and Scholz suggested, i.e., by getting rid of the false idea that possessing fewer false beliefs in a domain than laypeople do is a necessary condition for one to be an expert. Then, I argued against Goldman that his threshold condition (or, non-comparative clause) unjustifiably constrains the attribution of expertise to a "God's-eye" perspective. This requirement is unnecessary if we simply want experts to exert a novice-oriented function, that is, to provide laypeople with accurate information they lack. Finally, I explained why an understanding-linked possession condition fares better than a truth-linked condition when it comes to fulfilling NO-CAP. If I am right about Denny's case, the proposed analysis shed also light on the fact that NO-CAP appears to be unsatisfactory as a definition of the expert's function, thereby anticipating what I shall argue in the next sections of this chapter.

As a concluding remark of this section, I shall briefly discuss in which respects the novice-oriented account of an expert differs from Jäger's recent account of a *Socratic epistemic authority* (2016). On Jäger's view, a Socratic authority is "someone who not only succeeds more often in attaining the truth [than the subject does], but who also is able to foster the subject's overall insight into the problem under consideration" (178-9). These accounts have two main elements in common: first, they assume that these kinds of epistemically superior subjects have a novice-oriented function to fulfill; second, they measure the epistemic superiority of both experts and Socratic authorities by the level of their respective understanding of a domain.

However, the accounts also differ in two important respects: First, experts need to be epistemically superior to most people in a domain, whereas Socratic authorities merely need to be epistemically superior to their interlocutors. Second, the novice-oriented account does not restrict the strategies an expert can use to help novices improve their epistemic position, while Socratic authorities must perform their function in a very peculiar way, namely by using their maieutic ability to ask questions that guide the interlocutor to understand things for themselves.[20] I shall return to these differences in the next section to discuss important features of the service conception of epistemic authority, which Jäger's account belongs to. For now, it

his Goodmanian idea of "symptoms" and the standard notion of "necessary condition" should be. Thus, I limit myself to acknowledging his similar intuitions, and stick to my project of providing a compelling account of expertise through conceptual analysis.

[20] I shall offer further considerations on Socratic authorities and the maieutic abilities in §3.3.3.

suffices to say that despite their commonalities, we have reasons to keep Goldman's notion of an expert and Jäger's notion of a Socratic authority distinct.

2.4 Against the Novice-Oriented Account of an Expert

Despite the attempt to provide a compelling novice-oriented account of an expert made in the previous section, I now want to show that this view is liable to at least two objections that make it far less attractive than it might initially seem. Specifically, the objections show that the novice-oriented account fails to comply with Goldman's functionalist approach and, in particular, with the methodological assumption introduced in §2.2. First, I shall argue that neither the understanding-linked condition nor Goldman's truth-linked condition suffices to determine what it takes to have the capacity to help laypeople solve problems in a given domain, as some further condition is required. Second, NO-CAP is too weak as a functional definition of expertise, in that it attributes to experts a role that other kinds of epistemic subjects—whom we would not be willing to consider experts in general—can easily satisfy.

Let us consider the first problem affecting the novice-oriented account, namely that fulfilling U-LC is not sufficient for one to satisfy NO-CAP. Indeed, there may well be subjects who have much better understanding than most people in a domain and nonetheless lack the capacity to help others improve their epistemic position in that domain. Establishing exactly which dispositions ensure that one has the capacity to help novices in some domain appears to be a very difficult task, one that probably involves empirical considerations too. For our purposes, it suffices to say that this capacity involves at least the following set of intellectual virtues:

> *Novice-oriented abilities*: virtues that allow an epistemic subject to properly address a layperson's epistemic dependency on them (e.g., sensitivity to the novice's needs, intellectual generosity, intellectual empathy, sensitivity to the novice's epistemic resources, and maieutic ability).

It should be evident that we do not get these abilities for free simply by having better understanding in a domain than most people do. As a matter of fact, there can be several ways in which a subject S who fulfills U-LC can fail to display novice-oriented abilities.

First, S might be unable to figure out what a layperson does not understand about their theories, claims, or scientific findings, which S takes to be extremely clear, thereby displaying insensitivity to the layperson's epistemic needs. Second, S might be unable to tailor their answer to the novice's questions: for instance, S might provide overly complex information for the novice to understand; alternatively, S might underestimate the novice's epistemic resources and offer overly simplistic information, thereby failing to satisfy their epistemic needs. Finally, S might be able to figure out the layperson's needs but lack the ability to ask those questions that guide them to understand things for themselves. In this particular situation, S displays a lack of maieutic ability and intellectual empathy with the novice's epistemic situation—i.e., with their level of understanding of the domain, with how their understanding could be broadened, deepened, or made more significant.

This list is not supposed to include every possible way in which a subject who has better understanding than most people can fail to be a virtuous *teacher*. It does show though that these scenarios are plausible, and therefore puts pressure on those who endorse a novice-oriented account of an expert. If we want to find people who can fulfill NO-CAP, we cannot simply look for individuals who have great understanding of a domain. These subjects need, additionally, to have novice-oriented abilities.

Before considering the second problem with the novice-oriented account, it is worth noticing that the same issue would arise if we stuck to the most plausible version of Goldman's veritistic condition, namely T-LCa. If having better understanding than most people do in a domain is not sufficient to identify who can fulfill NO-CAP, *a fortiori* having more true beliefs than most people do cannot do any better. This is because T-LCa can be met by subjects who do not even grasp the relationships between different parts of a body of information, which is a fundamental feature that allows one to address others' questions and help them improve their epistemic position in a given field. Thus, we can conclude that no matter whether we endorse a truth-linked or an understanding-linked possession condition of expertise, further dispositional requirements are needed for one to satisfy NO-CAP.

The second problem with the novice-oriented account is that it offers an overly weak functional definition of what it takes for one to be an expert. All that is required for one to satisfy NO-CAP is that they are better epistemically positioned than *the interlocutor* is in that domain. The recent debate on the notion of *epistemic authority* which I will analyze in detail in the next chapter can help us clarify this problem. For the purposes of this chapter, suffice it to say that on the service conception of epistemic authority (Jäger 2016; Raz 1986, 2006; Zagzebski 2012), a subject is an epistemic authority for another insofar as they can help their

interlocutors achieve some epistemic goal in a given domain through their superior knowledge and/or understanding to the interlocutors. Thus, the function of epistemic authorities on the service conception is just a slightly revised version of Goldman's NO-CAP.

However, a relevant difference between the two accounts emerges when we compare their respective possession conditions. Setting aside for the time being the details of the various versions of the service conception of epistemic authority, let us assume that a subject S is an epistemic authority for a subject S* in domain D insofar as S (i) is more reliable than S* in D and (ii) makes use of their novice-oriented abilities to answer S*'s questions in D.[21] As I briefly anticipated in the previous section while discussing Jäger's view, the possession condition of epistemic authorities illustrates that the novice-oriented function can be fulfilled by subjects who are merely better epistemically positioned than their interlocutors but may not be better positioned than most people are in D, as U-LC or T-LCa require.

For example, a grandmother—let's call her Loren—can help her grandson acquire knowledge or understanding of how fish breathe simply by having some vague and basic understanding of zoology while he knows nothing about that. Supposing Loren also has some novice-oriented abilities through which she helps her grandson improve his understanding of ichthyology, she proves to be an epistemic authority for the grandson and, at same time, she fulfills NO-CAP.[22] But it should be evident that neither of them is better epistemically positioned than most people are in the respective domains. This is a problem for the novice-oriented account because it shows that rather ordinary epistemic subjects, simply by being in an epistemically privileged position with respect to their interlocutors, can satisfy its functional definition, which should capture a necessary and sufficient capacity for one to be an expert.

More will be said in §3 about the distinction between experts and epistemic authorities, but this quick comparison should be enough to show that they are two different kinds of epistemically proficient subjects and hence to create trouble for Goldman's novice-oriented approach. If he wants to hold firm that NO-CAP is a functional definition of an expert and to extract possession conditions by reference to it—as his methodological assumption requires—

[21] This preliminary definition is meant to capture the core tenets of the service conception of epistemic authority, which I shall motivate and defend in detail in chapter 3. For the moment, we can work with the above definition provided that we keep in mind two important features of the view (on which I shall return in the next chapter). First, Zagzebski's original account is grounded in the notion of "epistemic conscientiousness" rather than in "epistemic reliability". Second, I work under the assumption that the notion of epistemic authority requires a minimum threshold meant to prevent cases in which a reckless person is an authority for an even more reckless interlocutor. Arguably, though, the reliability threshold for epistemic authority is lower than the threshold for expertise.

[22] Notice that the account of an epistemic authority I will defend in §3 is more concessive than Jäger's in that it contemplates the possibility that an authority fulfills the novice-oriented function through their maieutic ability but also admits further strategies: for example, an authority can simply teach what they know or present their interlocutor with explanatory stories that help them acquire understanding in some domain.

he is forced to give up on T-LCa or U-LC and endorse weaker requirements. For NO-CAP merely requires that one be better epistemically positioned than the interlocutor is—yet not than most people are in a domain, as U-LC demands. But that is no less problematic, because intuitively it seems wrong to contend that epistemic authorities such as Loren possess what it takes to be an expert simply by being better epistemically positioned than their interlocutors in a given domain. These considerations allow us to conclude that the novice-oriented account does not comply with Goldman's functionalist approach. But, as I have anticipated, all is not lost: in the next section, I shall argue that the research-oriented view is not liable to the same objections.

2.5 Toward a Research-Oriented Account of an Expert

A broader question one might want to ask about the novice-oriented account is what are the legitimate and plausible ways an expert can fulfill their novice-oriented function. Surely teachers fulfill this role, as I argued in the previous section, but on a more charitable reading of NO-CAP it might seem as though also *researchers*—that is, figures such as Marie Curie or Rita Levi Montalcini—should be included. As a matter of fact, researchers' contribution to the progress of a discipline, now or in the future, will allow themselves or others to answer the novices' questions and help an entire community enhance collective knowledge and understanding in that field.

On closer inspection, though, this charitable reading turns into a further problem for the novice-oriented approach. Consider the case of a chemistry researcher—let's call her Shyler—who, despite her contribution to the progress of her field through her discoveries and publications, is completely deprived of novice-oriented abilities and social skills. Shyler's social interactions are limited to what is necessary for keeping her research position at some institution; she conducts her work in isolation as much as possible. I take it that the very fact that she made discoveries in chemistry intuitively suffices to grant that Shyler is an expert in that domain, no matter her social limitations. If so, then the novice-oriented abilities and the capacity to help laypeople acquire knowledge or understanding in a domain become less than necessary features of experts. To accommodate a case like this, the novice-oriented account would have to pay a huge cost, in that it should water down its novice-oriented spirit to such an extent that it is not anymore clear that Shyler and Loren have a specific capacity in common.

These considerations allow us to bring the research-oriented account back on stage. As I shall argue in the rest of this section, the research-oriented view avoids these unwelcome results and complies with Goldman's methodological assumption better than the novice-oriented account does, in that it is not liable to the objections raised in the previous section against the novice-oriented account. However, before developing the arguments in favor of the research-oriented view, the account has to be refined in light of the problems affecting the truth-linked condition discussed in §3.

In the original version of the research-oriented account, expertise is defined as someone's capacity to give a contribution to the progress of some domain (RO-CAP), and the possession of more true beliefs and/or fewer false beliefs (T-LC2) is considered as a necessary and sufficient condition for one to fulfill RO-CAP. In §2.3.1, I offered reasons to replace T-LC2 with T-LCa. However, it seems clear that replacing T-LCa with the understanding-linked condition (U-LC) has better prospects for success for there are many ways in which someone can possess more true beliefs than most people do in a domain without having the capacity to contribute to the progress of that domain. To consider one of them, let us go back to EXPERT BY HEART (§3.3). The fact that Denny fulfills T-LCa does not put him in a position to satisfy RO-CAP—that is, does not ensure that he has the capacity to answer new questions arising within a domain. After all, all he can do is reply to specific questions about physiology whose answers he learnt by heart, but he has no idea of the current challenges of this field and how to solve them, due at least in part to a lack of understanding.

As I anticipated in §2.2, someone might want to resist this argument, holding that if the propositions pertaining to a domain of expertise D include propositions about the processes, methods, and skills the use of which is likely to yield true beliefs in D, then satisfying T-LCa will typically result in the capacity to generate new knowledge in D, at least in non-empirical domains. I think this move cannot settle the issue. We could think of a slightly revised version of EXPERT BY HEART in which Denny helps his brother with a difficult algebra exam, and specify that the information Denny learns by heart includes propositions about the processes, methods, and necessary skills one needs to deploy in the field. But it is hard to see how this could allow him to contribute to the progress of the research in algebra. The fact that he has lots of true beliefs about that domain and about its research methodologies does not ensure that he understands how these pieces of information connect to each other and has the capacity to use them to conduct research in algebra. Thus, I think we have good reasons to abandon T-LCa again and endorse U-LC as the possession condition on the research-oriented account.

The rest of the section is devoted to arguing that this view has better prospects to succeed as a functionalist account of expertise. The first, general, task is to show that the research-oriented account, unlike the novice-oriented one, can accommodate our intuition that figures such as Marie Curie but also Shyler are experts while Loren is not. This view accounts well for such an intuition: specifically, it grants expertise to the former as they satisfy the research-oriented function, while the latter do not and therefore lack expertise. Indeed, Loren has the capacity to address the lay interlocutor's question by virtue of having better understanding than the interlocutor in the respective domains, yet both lack the capacity to answer new questions arising within them. As a matter of fact, Loren only has a rough understanding of some basics of ichthyology. That is good news for the research-oriented view because it shows that this account, unlike the novice-oriented one, is not committed to grant epistemic authorities the status of experts.

The second task is to verify whether the research-oriented account complies with Goldman's methodological assumption. To do that, we need to have positive reasons in favor of the idea that U-LC is a necessary and sufficient characterization of what it takes for one to fulfill RO-CAP. Making a case for the *necessity* of U-LC is a relatively easy task. Surely, one of the reasons why Rita Levi Montalcini managed to contribute to the progress of medicine with her work in neurobiology is that she had a much better understanding of human physiology, as well as several other branches of medicine, than most people have. U-LC is a necessary feature of those who fulfill RO-CAP because this capacity involves at least a solid grasp of the relationships that link various elements of a domain, previous issues that other people had to solve within the domain, and questions that need to be addressed to advance our competence in the domain. But only people who have great understanding of a discipline can grasp these elements.

Yet it is far from clear that U-LC can be a *sufficient* requirement for individuating who satisfies RO-CAP, as there may be subjects who, despite having much better understanding than most people do in a domain, still lack the capacity to contribute to its progress. Suppose, for instance, that an outstanding scholar who has published dozens of important articles in her field is the victim of a terrible car accident, which impairs her ability to do research and acquire new information, without compromising her memory (see Coady 2012: 30). In such circumstances, the scholar still satisfies U-LC but since she is no longer able to keep updated on the progress of her discipline, she loses the capacity to contribute to the research in the field, thereby failing to satisfy RO-CAP.

Besides the case of the impaired researcher, a more general reason why a subject who has great understanding in some field might not satisfy RO-CAP is that they may lack important abilities that are necessary for contributing to the research in any field. As in the case of the novice-oriented account, we cannot arrive at a conclusive set of necessary and sufficient dispositions, but RO-CAP seems to involve at least the following set of intellectual virtues:

> *Research-oriented abilities*: virtues that allow an expert or authority to exploit their fund of knowledge and understanding to find and face new problems in their field of expertise (e.g., thoroughness, intellectual perseverance, intellectual courage, self-scrutiny, intellectual creativity, open-mindedness, intellectual curiosity, and autonomy).

Coady would disagree with this analysis and argue that having great understanding of a domain already shows that one possesses research-oriented abilities and therefore conclude that these abilities add no extra requirement to the possession condition (2012: 29). On closer inspection, this claim seems wrong. As a matter of fact, someone may have great understanding of a domain because they are intellectually virtuous and dedicated subjects, or because they just have intellectual talent. But a talented subject could lack the capacity to address new questions arising in a domain because, for example, they lack the curiosity that makes one ask good questions about the problems of the domain, the creativity that makes one find possible answers to those problems, the thoroughness that is necessary to notice opportunities for progress in small details, the perseverance that is necessary to carry on with a project despite objective obstacles, the courage that allows one to defend their view in the face of potential harm, or the open-mindedness that allows one to consider alternative standpoints.[23] Thus, someone might possess great understanding of some field and nonetheless fail to contribute to the progress of that field because of a lack of research-oriented abilities, which therefore appear to be necessary for a subject who displays great understanding to satisfy the functional definition of expertise on the research-oriented account.

In conclusion, this analysis has shown that the research-oriented account, unlike the novice-oriented account, complies with Goldman's functionalist approach, as it provides (i) a functional definition of expertise that other kinds of epistemic subjects—namely, epistemic authorities—are unable to fulfill (i.e., RO-CAP), and (ii) a possession condition (i.e., U-LC plus research-oriented abilities) that is at least necessary for individuating those subjects who can

[23] Further considerations on the distinction between intellectual talent and intellectual virtues are offered in §7.3.1.

satisfy the functional definition. Furthermore, the research-oriented account accommodates several intuitive cases that threaten the plausibility of the novice-oriented approach. Specifically, it allows us to contend that Shyler is an expert because she has the capacity to further the progress of chemistry through her great understanding of the domain and the research-oriented abilities she employs to conduct research. In contrast, it entails that Loren is not an expert, because she can merely help her interlocutor understand something better in a given domain but lack both better understanding than most people have in that domain and the abilities to contribute to its progress.

The arguments provided in these sections allow us to conclude that the research-oriented account better explains what it takes for one to be an expert than the rival, novice-oriented approach. The fact that the set of research-oriented abilities largely differs from that of novice-oriented ones helps us explain why neither functional definition of the two competing accounts can be taken to include the other. Just as epistemic authorities have the capacity to help others without necessarily possessing the ability to make contributions in any domain, so experts can contribute to the progress of their domain without necessarily having the capacity to help their interlocutors acquire knowledge or understanding in that domain.

2.6 Conclusion

This chapter surveyed two major approaches to the notion of the cognitive expert, based on Goldman's idea that a compelling definition of an expert needs to reflect the function of expertise that these subjects fulfill in the epistemic community. The main goal of the chapter was to start addressing the Conceptual Question that grounds this thesis by showing that a research-oriented account complies with Goldman's functionalist approach, while a novice-oriented account fails to do so. To pursue this aim, in section 2.2 I introduced the basic features of both views and showed how their functional definitions are meant to fall in line with their respective possession condition, as required by Goldman's methodological assumption. In section 2.3, I refined the possession conditions of both accounts to make sure we were working with the best versions of them. In section 2.4, I developed an argument against the novice-oriented account by showing that its functional definition commits us to grant epistemic authorities the status of experts, against our broad intuitions about expertise. Finally, in section 2.5, I argued that the research-oriented account is not liable to the problems affecting the rival approach and that it complies with Goldman's methodological assumption.

As a final remark, I want to go back to the distinction between experts and epistemic authorities introduced in §2.4. It may have been noticed that in a few passages I characterized the difference between epistemic authorities and experts by referring to *teachers*, i.e., those who are trained to help novices acquire knowledge and understanding of a given domain, and *researchers*, i.e., those who are trained to foster the epistemic progress of a discipline. In the ideal scenario, we would rather be taught by researchers: the more this happens, the better it is for everyone. For, when an expert possesses novice-oriented abilities, laypeople can acquire fresh knowledge and understanding from someone who works first-hand in the domain. This is beneficial for novices not only because it minimizes their risk of getting improper or outdated information, but also because an expert can direct them towards the right kind of issues that they need to explore if they want to increase their understanding of D.[24]

As I tried to argue in sections 2.4 and 2.5 though, we cannot take it for granted that experts possess novice-oriented abilities. Those who work in education know how important it is that teachers be trained to perform their task effectively. It might well be the case that great researchers prove to be less effective as teachers than trained teachers who lack research-oriented abilities. An important consequence of the argument of this chapter for real-life scenarios and educational policies, then, is that we would rather have experts working on advances in their fields than spending lots of time teaching novices when they lack the capacity to do so. For, if my argument hits the target, it should be clear that experts can do a better service to the whole society by fulfilling their research-oriented function than by working as—or, trying to become—epistemic authorities.

[24] More on the idea of experts and epistemic authorities as advisors can be found in Elga (2007), Jäger (2016), and Lackey (2018).

Chapter 3

An Alternative Account of Epistemic Authority: Limited-Preemptionism and the Role of Intellectual Virtues

Abstract. According to a recent account of epistemic authority proposed by Linda Zagzebski (2012), it is rational for laypersons to believe on authority when they conscientiously judge that the authority is more likely to form true beliefs and avoid false ones than they are in some domain. Christoph Jäger (2016) and Jennifer Lackey (2018) have recently raised several objections to her view. I argue that both theories fail to adequately capture what epistemic authority is, and I offer an alternative account grounded in the abilities that different kinds of authorities are required to possess.

3.1 Introduction

A recent book by Linda Zagzebski, *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief* (2012), offers an innovative account of authority in the epistemic realm. Her account focuses on the necessary conditions that legitimate believing on authority or, in other words, on the question: whom should we believe? (Coady 2012: 27). Christoph Jäger (2016) and Jennifer Lackey (2018), among others,[25] raise several objections to Zagzebski's view and endorse alternative theories of epistemic authority. In this chapter, I suggest that all these views are incomplete: my twofold aim is to solve the issues arising from these works and to provide a comprehensive account of epistemic authority that includes not only the expert and the authority of (true) belief, but also the authority of understanding.

By doing so, the chapter purports to address the Conceptual Question and the Normative Question that lie at the grounds of this thesis. As regards the former question, the chapter contributes to explaining what it takes for one to be epistemically superior to another by distinguishing the notion and the function of epistemic authorities from those of experts. As

[25] Further critiques to Zagzebski's preemptionist theory of epistemic authority have been raised in Dormandy (2018), Gasparov (2017), Grundmann (forthcoming), and Ward (2017).

regards the latter, the chapter defends what I call a limited-preemptionist view of epistemic authority, that is, the idea that there are situations in which the most rational attitude that novices could—and, in fact, should—adopt before expert testimony is that of deferring to their epistemic authority.

In section 3.2, I introduce Zagzebski's innovative definition of epistemic authority, which clearly distinguishes authorities from experts in the epistemic domain. In section 3.3, I present the *Preemption Thesis*, a crucial condition of Zagzebski's account of epistemic authority, and discuss Jäger's fundamental objections to Zagzebski's theory of authority. According to Jäger, Zagzebski's theory fails to account for the importance of epistemic authority in domains like philosophy and scientific research, where understanding is a more important epistemic aim than having true beliefs. In section 3.4, I offer three compelling considerations for going beyond Zagzebski and Jäger's accounts. In section 3.5, I discuss Lackey's objection to Zagzebski's theory and her own alternative view of epistemic authority. Building on these remarks, in section 3.6 I develop a pluralistic and virtue-based account of the epistemic authority that includes the three main kinds of authority in the epistemic domain—namely, experts, authorities of belief, and authorities of understanding.

3.2 Experts and Epistemic Authorities

As anticipated in §1.1, the issue of defining epistemic authority can be approached from two different perspectives: on the one hand, authority refers to "the normative power possessed by the bearer of authority" (Zagzebski 2012: 103), as most epistemologists have usually understood it. On the other hand—following Zagzebski's classification[26]—authority can indicate "the person or institution that has authority" (103).[27] Here Zagzebski sheds light on a deeply-rooted confusion concerning the former definition, which has led most to consider epistemic authority merely as a property of those who are experts in some particular field, and hence to refer to the very same persons as "authorities" or "experts" indistinctly.

Conversely, Zagzebski approaches this notion from the latter point of view, defining epistemic authority as follows:

[26] In this chapter I use "EA" and, later, "SA" when I talk about "epistemic authorities" and "Socratic authorities" as persons. Furthermore, I call "S" the layperson or the novice epistemic subject and, for the sake of simplicity, I consider novices as males, but experts, EAs, and SAs as females.

[27] Further considerations on the distinction between *being an* authority and *having* authority can be found in Raz (1979: 19-20; and 2006: 1034).

> *Epistemic Authority.* Someone who does what I would do if I were more conscientious or better than I am at satisfying the aim of conscientiousness—getting the truth. (109)

Since this is not a commonly shared definition among epistemologists, it will be wise to clarify the role played by "[epistemic] conscientiousness". Zagzebski defines it thus:

> *Epistemic Conscientiousness.* Reflective disposition to use our faculties as best we can for obtaining the truth.

As a disposition, conscientiousness is matter of degrees and goes from the basic level of self-awareness we have most of the time (i.e., awareness that we are agents with abilities) to a more significant level in which we self-reflectively aim at getting the truth in various ways (48-49). Thus, the epistemic authority is a subject who is more epistemically conscientious than I am, and hence is more likely than I am to obtain the truth in a particular field.[28]

I suggest that the notion of an authority differs in three main respects from the notion of a cognitive expert that I have offered in the previous chapter. The first difference pertains to the nature of the relationship between the epistemic superior and laypeople. Zagzebski's definitions make the notion of epistemic authority clearly *comparative* or *subject-dependent*, since she conceives it as a three-place relation between a given subject (or set of subjects) S, another subject (or set of subjects) S*, and some doxastic domain D.[29] Suppose you are used to picking mushrooms in the countryside: it might easily be the case that someone is an EA for me with respect to recognizing different kinds of mushrooms, while she is not an EA for you in that particular domain. Analogously, someone might be an EA for me with respect to distinguishing mushrooms, whereas she is not an EA for me when it comes to philosophy of religion.

In contrast, as we have seen in §2.3, the notion of an expert possesses a *non-comparative* or *subject-independent* dimension. Here is a simple way to recast the first difference between the subject-independent notion of an expert and the subject-dependent notion of epistemic

[28] It is wise to mention that, following Zagzebski, I work under the assumption that S is conscientious as well, that is conscientious enough to realize that he has evidence that some individuals are more conscientious than he is in a given domain (see Zagzebski 2012: 57). This consideration sets a threshold that is meant to prevent cases in which a reckless person is an EA for an even more reckless interlocutor.

[29] Jäger argues that epistemic authority is also time-relative: "[s]ince our epistemic skills and virtues and the amount of available evidence and information about a given topic vary over time, S* may be an authority for S at a given time t_0 but no longer play that role at some later time t_1, or vice versa" (2015: 4).

authority. Granted that Mourinho is an expert in football tactics, he is an EA *for* all other amateur football players and coaches. In contrast, he might not be an EA *for* his mentor Louis Van Gaal or *for* Carlo Ancelotti—who won the Champions League more times than Mourinho—as they might easily be as epistemically conscientious as Mourinho.

The second difference between the two notions is explicitly suggested by Zagzebski when she claims that "the expert is an authority only in a very weak sense, since the expert and the layperson who defer to her may have no relationship with each other" (2012: 5). Surprisingly enough, Zagzebski does not provide an argument in support of this claim. What we can extract from this quote is that a further condition for one to be an epistemic authority—in the full or strong sense—is that one shares a relationship with the interlocutor. Why so?

I hold that having a personal relationship with someone epistemically superior to us in some domain is a necessary condition for her to be able to acknowledge our epistemic dependence and needs. As I will argue in the rest of the chapter, the role of epistemic authorities can go far beyond the mere transmission of true beliefs to the novice. Indeed, I will defend the idea that someone is an authority for another insofar as she responds to (some of) his epistemic needs, which may vary, depending on the circumstances, from getting knowledge to understanding a subject matter, and from receiving what he needs through the authority's testimony to achieving epistemic goals on his own thanks to the authority's guidance.[30] But the authority cannot be sensitive to the novice's needs if she does not have a relationship with him: thus, in absence of this condition, the authority might fail to provide what the novice is looking for.

In contrast, having a relationship with the novice seems unnecessary for one to be an expert, as neither the understanding-linked condition (U-LC) nor RO-CAP require to. As we have seen in §2.5, the mere fact that Shyler, the scholar who lacks novice-oriented abilities and social skills, barely lacks any interaction with her colleagues and has no students fails to undermine her expertise. Nonetheless, it should be evident that she cannot be an epistemic authority in Zagzebski's strong sense, due to her lack of a relationship with laypeople. On the other hand, as I pointed out with the example of the authority in recognizing mushrooms, it might be the case that someone is an EA for me in some field, without necessarily being a socially recognized expert in that field.

The last difference between the expert and the epistemic authority concerns the intellectual virtues each is required to possess. In the previous chapter, I have shown that research-oriented

[30] This notion of the epistemic authority is in line with the Razian model of the "service conception" of the authority, according to which a subject A is an authority for a subject B insofar as A is doing something, e.g., issuing binding directives, in the service of B (see Raz 2006).

abilities are fundamental to experts, yet not to EAs: in fact, someone might be an EA for me in North-European contemporary literature simply by having attended an undergraduate course on that topic, given that I do not know anything whatsoever about Baltic writers. Such a student could inform me about the most important writers and their work, yet be unable to generate new knowledge on Baltic contemporary literature, thereby lacking any expertise on that topic.

In section 3.6, I will distinguish between various kinds of authority depending on the abilities each needs to possess. At this stage, the analysis of the notion of an expert, combined with Zagzebski's considerations on epistemic authorities, allow us to conclude that we should not treat "epistemic authority" and "expert" as overlapping notions. The former is subject-dependent, whereas the latter is mainly subject-independent. Moreover, the former, unlike the latter, requires an ongoing relationship between the layperson and EA. Finally, the latter, unlike the former, demands that the expert be able to generate new knowledge in her field of expertise.

Following the approach endorsed in the previous chapter, I shall again offer a preliminary definition of an epistemic authority—namely one based on Zagzebski's theory—and then provide reasons throughout the chapter for adopting a more promising definition. For now, we should work with the following definition:

> *Epistemic Authority (EA)*. An epistemic subject A is an EA for another subject S in domain D iff:
> (i) A is *more epistemically conscientious* than S in D;
> (ii) There is an ongoing relationship between A and S.

3.3 Zagzebski's Preemptionist Account and Jäger's Objections

3.3.1 Zagzebski's Account of the Authority of Belief

Zagzebski's account of epistemic authority is based on two fundamental theses she derives from Joseph Raz (1988): the Preemption Thesis and the Justification Thesis 1 for the Authority of Belief.[31] Both are meant to provide justification for a seemingly counter-intuitive action, such

[31] For the Razian version of the Preemption Thesis see Raz (1988: 42, 57-59; and 2006: 1019, 1024); Zagzebski's Justification Thesis 1 for the Authority of Belief is grounded in Raz's Normal Justification Thesis (1988: 53). Arnon Keren offers an interesting discussion of the Preemption Thesis in his (2014a).

as believing on authority. The former is "a thesis about what it means to take someone as an epistemic authority" (Zagzebski 2014: 176) and can be defined as follows:

> *Preemption Thesis for epistemic authority:*
> The fact that the authority [EA] has a belief *p* is a reason for me to believe *p* that replaces my other reasons relevant to believing *p* and is not simply added to them. (Zagzebski 2012: 107)

Thus, according to the Preemption Thesis, believing on authority means that if I consider someone to be an epistemic authority for me, then I am not only in the position to defer to her; rather, I am epistemically compelled to adopt her reasons for belief as the decisive ones to settle the matter. I am not required to dismiss or completely disregard my own reasons for believing *p* as long as I let the authority's reasons be those for why now I believe *p*.[32] However, although the Preemption Thesis accounts for the epistemic attitude I am rationally supposed to adopt, it says nothing about what justifies my adopting EA's reasons for belief.

Zagzebski's latter thesis explains "the normal way to show that [E]A is an authority for S" (2014: 177) as follows:

> *Justification Thesis 1 for the Authority of Belief (JAB1).*
> The authority of another person's belief for me is justified by my conscientious judgment that I am more likely to form a true belief and avoid a false belief if I believe what the authority believes than if I try to figure out what to believe myself. (Zagzebski 2012: 110)

According to JAB1, what allows me to consider someone an EA is my reflective judgment that she is more conscientious than I am, which in turn makes it rational for me to believe what she believes if I want to increase my chances of getting the truth.

Let us consider an ordinary example:

> RED CARDINAL DISAGREEMENT CASE. Suppose I form the perceptual belief that there is a chaffinch on the tree in front of me, but I am not completely confident of

[32] See sections 3.4.2 and 3.5 for further clarifications on the functioning of preemptive reasons. See also Jäger (2016) and Zagzebski (2016).

it. Suppose, then, that I ask an ornithologist who is passing by and she says that the bird is a red cardinal (*p*).

According to Zagzebski, once I reflectively judge through JAB1 that I am more likely to form a true belief if I trust the ornithologist and believe that *p* than if I try to figure out what to believe myself, I have a compelling reason to consider her an EA and to adopt her belief, as demanded by the Preemption Thesis.

3.3.2 Jäger and Zagzebski's Dispute on the Preemption Thesis

As mentioned above, there is debate over whether believing on authority requires preempting. Jäger, among others, raises several concerns against the Preemption Thesis: here I focus on what he calls "the problem of unhinging proper bases" (see 2016: 173-ss.). This problem does not directly pertain to cases of initial disagreement between S and EA, such as the one considered in the previous example.[33] Rather, it concerns cases of partial or full agreement between S and EA that, according to him, Zagzebski's account is not in a position to accommodate.[34]

In order to illustrate Jäger's first objection, consider the following example:

> RED CARDINAL AGREEMENT CASE.[35] Suppose that I form the perceptual belief that there is a red cardinal on the tree and that I am not entirely confident of it. Suppose, then, that I ask the passer-by ornithologist whom I consider an EA and she confirms my belief that *p*.

The Preemption Thesis, Jäger notices, requires that I stop considering my visual experience as the reason for which I believe *p* and instead preempt; but in such a situation where I partially (or fully) agree with EA, (i) it might not be psychologically possible for me to disregard my evidence and (ii) I need not do so. Clearly, the fact that EA shares my belief is additional supporting evidence for my true belief, which is (partially) justified by my visual experience.

[33] Notice that Jäger's problem of unhinging proper bases is part of a more detailed critique of Zagzebski's account of epistemic authority, where he individuates two further problems with preemption—namely, the problem of competing authorities and the switching problem (see 2016).
[34] Where a case of partial agreement is a situation in which both S and EA have the same belief, yet based on different epistemic reasons.
[35] The example is adapted from Jäger (2016).

Grounding my belief in both my own perceptual evidence and the authority's judgment seems epistemically better off than simply supporting such a belief with uncertain perceptual reasons. Nonetheless, there is no reason why I should set aside my own evidence, since it revealed itself as accurate.

Zagzebski's reply to Jäger's objection aims at deflating the relevance of the subject's own reasons compared with the authority's. She argues that there is nothing I can see about the bird that the ornithologist has not already noticed. Thus, it would be illusory to think that adding my own reasons to the ornithologist's would provide a better justification for believing that *p* (see Zagzebski 2016: 190). Furthermore, Jäger's remark that it might not be psychologically possible to preempt in cases of agreement like this should not affect Zagzebski's account, which is meant to provide a rational way to justify believing on authority, rather than determining whether or not it is possible to preempt in such and such a situation. What matters to her is that preemption assures the best track record in terms of getting the truth.

Without entering the debate too deeply and straying too far from my purposes, there is a compelling reason why Zagzebski's reply does not make her argument safe from Jäger's objection. Consider the following modified version of the RED CARDINAL AGREEMENT CASE, where the authority is not a socially recognized expert, e.g., an ornithologist, but just someone who is better epistemically placed than I am.

> RED CARDINAL SISTER AGREEMENT CASE. Suppose I believe there is a red cardinal on the tree due to hearing a particular song that I ascribe to red cardinals. Then, suppose I ask my sister whether the bird on the tree is a red cardinal, since I have just lost my glasses and I am not entirely confident of my belief. She has no idea what the red cardinal's song sounds like, but she reports that the bird is a red cardinal (*p*) on the basis of strong visual evidence that the brilliant red bird with a red mask on the tree is a red cardinal.

If we assume that there are circumstances, as in the RED CARDINAL SISTER AGREEMENT CASE, in which it is not necessary for one to be an EA to already be aware of possible reasons in support of *p*, then preempting might not be the most rational thing to do for S. Let us grant that deferring to my sister, i.e., the epistemic authority here, is rational, since she is better epistemically placed than I am. Nevertheless, despite my nearsightedness, when I asked for her help, I already had some reasons in support of my belief that were not part of her justification for *p*—namely, auditory evidence concerning the red cardinal's song. These perceptual reasons

enforce both my justification for *p* and my confidence that *p* is the case. Therefore, preempting might still lead to the best track record in getting the truth, yet considering my sister's testimony as a further evidential reason for my belief that *p* would give me a stronger justification for *p* than she has. Hence, contra Zagzebski, it would be irrational for me to stop also basing my belief on my auditory evidence and just preempt.

3.3.3 Jäger's Account of Socratic Authority

Jäger's second objection to Zagzebski's account of epistemic authority emerges in cases of initial disagreement between S and EA within important domains like philosophy, scientific research, politics, and religion, or in those ordinary situations where gaining understanding is a more fundamental goal than merely having true beliefs. In such domains, the Preemption Thesis fails to capture what epistemic authority actually is, because the layperson S who disagrees with EA would not be willing to simply switch beliefs, adopt EA's, and improve his track record of true beliefs. What S wants is to enhance his understanding of some subject matter, that is, "grasping systematic connections among elements of a complex whole, or gaining insight into certain relations between items within a larger body of information" (Jäger 2016: 180).

Let us clarify the objection with an example: suppose I falsely believe that the recipe for tuna tartare requires that I cook the tuna in the oven for half an hour, but a chef, i.e., the epistemic authority for me here, tells me that I ought not to cook the tuna in the oven when preparing a tuna tartare. According to Zagzebski's account, once I conscientiously judge that the chef is more likely to get the truth than I am, I should just preempt my own reasons and adopt her belief. However, preempting would not give me any further understanding of the tuna tartare recipe, which is what I want to learn. I would expect an epistemic authority to improve my understanding by presenting further reasons in support of her belief. Merely knowing that I need not cook the tuna in the oven would not be of much help: acquiring more understanding would entail that I also learn, for instance, that tuna tartare is made with raw tuna, sliced into little strips or cubes, and served with different sauces or salads depending on the cook's preferences. This kind of expectation from an epistemic authority is something Zagzebski's theory does not explain, as she does not take a stand on whether her account of the authority of belief can also apply to understanding.[36]

[36] Moreover, Zagzebski holds that understanding cannot be transmitted through testimony, as I will show in section 3.4.2. However, it might be argued that Zagzebski's theory of epistemic authority is in a position to account for

In contrast, Jäger argues that we should consider authorities in the epistemic realm those who are concerned with fostering S's understanding rather than merely providing him with true beliefs. This is what he calls a Socratic authority:

> *Socratic authority (SA).*
> If someone is a Socratic authority for a subject in a given domain, he not only has a higher ratio of true to false beliefs in the domain than the subject does. He also displays superior methodological skills and insights which enable him properly to assess evidence, reasons, methods of thinking and investigation, and so on, and to communicate such insights to others. […] A Socratic authority, in other words, serves not only as a source of maximizing true beliefs at the object level, but also as a source of *understanding.* (Jäger 2016: 179)[37]

A comprehensive account of epistemic authority has to make room for Socratic authorities. The example of Socrates shows that the notion of epistemic authority not only serves as an efficient "litmus paper" for true beliefs, on Zagzebski's view; it also ought to account for the epistemic guidance that authorities offer, as suggested by Lackey (see §3.5). I will be accounting for the authority of understanding in the next section. Here, it is worth adding a final remark: it is wise to separate the negative part of Jäger's project, that is undermining the Preemption Thesis, and its positive proposal that we define authority in the epistemic realm in terms of Socratic authority. For admitting SA as a fundamental level of epistemic authority, as in fact I do, does not commit us to accepting Jäger's argument against the Preemption Thesis. Therefore, in what follows, I suggest that there are good reasons for both saving the Preemption Thesis and offering a theory of epistemic authority that is able to account for the authority of belief as well as for the authority of understanding.

---

the expectation that the chef explains the whole recipe to me, since EAs are, by definition, sensitive to S's needs. If so, then epistemic authorities à la Zagzebski have the potential to enhance S's understanding. In the next section, I will offer an argument for broadening Zagzebski's view in such a way that we can consider the notion of epistemic authority as including both authorities of belief and authorities of understanding.

[37] Some might think that Jäger's definition of Socratic authority still has a too veritistic flavor. Following Kvanvig's (2003) distinction between central and peripheral beliefs, we could argue that a Socratic authority need not have a higher ratio of true to false beliefs in the domain than the subject, provided the authority at least has a higher ratio or number of true *central* beliefs about the subject matter than the other individual. This remark sounds plausible to me. However, for the sake of simplicity I leave it aside here.

3.4 Going Beyond Zagzebski and Jäger

This section outlines the critical side of my project, which sets the ground for the alternative account I will offer in section §3.5. First of all, I will argue that Zagzebski's theory of epistemic authority could potentially explain how an authority can promote understanding, which is as much of a fundamental epistemic goal as having true beliefs.[38] Then, I will show that both Zagzebski and Jäger are wrong about preempting: on the one hand, Jäger demonstrates that there are several cases where Zagzebski's Preemption Thesis does not explain what epistemic authority is. On the other hand, pace Jäger, there are particular circumstances in which preempting can still be the most rational option for a layperson. Finally, I will defend the claim that Socratic authority constitutes a special model of epistemic authority, rather than an alternative account of epistemic authority.

3.4.1 The (Unexpressed) Potential of Zagzebski's Account

As anticipated, this section is devoted to defending the thesis that epistemic authorities need not limit the range of the epistemic goods that they promote to providing novices with mere knowledge. In the previous chapter, I have argued that possessing better understanding than most people do in a domain is a necessary condition for one to be an expert. Here I shall argue that epistemic authorities too can possess better understanding than their interlocutor and, in fact, that their function can well encompass fostering the novice's understanding of a given subject matter. More specifically, I shall argue in favor of the following claim:

> *Claim 1*. There is no valid reason to limit an account of epistemic authority to the authority of belief; indeed, it can be expanded so as to include the authority of understanding.

Consider the following example:

> HISTORY PROFESSOR CASE. Suppose a student S believes that George W. Bush sent US troops to Iraq to depose Saddam Hussein, and that during a history class, the

[38] See, for instance, Carter and Gordon (2014), Kvanvig (2003: 202-203) and Pritchard (2009: 39-40).

teacher says that Bush started fighting against Iraq also because Iraq possessed many oil wells.

Such an example of broader understanding, a very common one within education, clearly shows that the teacher's considerations, which were in fact absent from S's initial understanding of the issue, would improve his understanding of the reasons why the US attacked Iraq. Indeed, they allow him to figure out why starting that war was potentially profitable from an economic and financial point of view. Thus, the HISTORY PROFESSOR CASE clarifies why S can consider the professor an epistemic authority even though they share a true belief—namely, that "George W. Bush sent US troops to Iraq to depose Saddam Hussein"—and they (partially) agree on the causes of the Iraq War.

For what makes the professor an authority for the student is not her having more true beliefs than him, but her possessing a broader understanding and sharing it with him. This is why she counts as an authority of understanding, rather than an authority of belief. Based on this example, we can introduce a thesis for the authority of understanding that parallels Zagzebski's JAB1:

Justification Thesis 1 for the Authority of Understanding (JAU1).
The authority of another person's understanding of a subject matter *x* for me is justified by my conscientious judgment that I am more likely to gain a broader, deeper, or more significant understanding of *x* if I believe what the authority believes than if I try to figure out what to believe myself.

This thesis expands Zagzebski's account, since it suggests another way to establish that EA is an authority for S, i.e., showing that S's chances of improving his understanding of something increase if S trusts EA, instead of sticking to his own view. Notice, moreover, that the notion of understanding plays a different role in the case of epistemic authority than in the case of experts: as regards the former, for a subject to serve as an authority of understanding for her interlocutor it is merely required that her understanding of *x* is superior to the interlocutor's in one of the three main dimensions of understanding according to Elgin (see §2.3.3); in contrast, as regards the latter, one can be an expert about *x* if her understanding of *x* exceeds the understanding of most people in all three dimensions.

3.4.2 Weakening but Extending the Preemptionist View

Claim 1 only concerns the way in which we can justify broadening the notion of epistemic authority so as to include the authority of understanding. Whether or not the Preemption Thesis works for the authority of understanding is a different matter, which I explore with the following consideration:

> *Claim 2*. A more comprehensive account of epistemic authority, one that contemplates both the authority of belief and the authority of understanding, need not set aside the Preemption Thesis.

Claim 2 can be split into two different questions: (i) whether it is possible to maintain that the Preemption Thesis explains what epistemic authority is, and (ii) whether this thesis can apply to both the authority of belief and the authority of understanding. Jäger concedes that in several ordinary circumstances preempting is still the most rational thing to do.[39] Suppose I go to the doctor and ask for some cure for my cough, but she tells me that, instead, I need to cure my acid reflux, which is the real cause of the cough. In such a case, once I conscientiously judge that my doctor is an authority for me in this domain, it is rational to give up my belief and the epistemic reasons supporting it and preemptively adopt the doctor's belief.

However, as discussed in sections 3.3.2 and 3.3.3, Jäger holds that preempting is irrational both when S has epistemic reasons regarding whether *p* is the case and in those domains where understanding is a more important epistemic goal than acquiring true beliefs. I disagree with him on both claims. Concerning the former, I will defend a form of *limited-preemptionism*, in that I shall argue that it is rational for S to let EA's reasons preempt his own both when S disagrees with EA and when they agree but EA's reasons include and exceed all the reasons S has for believing *p*. As for the latter, I will demonstrate that a refined version of the Preemption Thesis is compatible with the notion of the authority of understanding. Thus, I endorse positive answers (partially, at least) to both (i) and (ii).

Consider (i). I suggest that the best way out of the Zagzebski-Jäger debate is to conditionalize the Preemption Thesis on particular, yet not-too-narrow, circumstances:

> *Preemption Thesis Weak for the Authority of Belief (PTW-B).*

[39] See Jäger (2016: 168).

> There are circumstances in which the fact that an EA has a belief *p* is a reason for me to believe *p* that replaces my other reasons relevant to believing *p* and is not simply added to them.

Weakening Zagzebski's Preemption Thesis in this particular way restricts its applicability to specific situations, thereby achieving two important objectives: first, the thesis accounts for those ordinary cases in which preempting is the most rational option for S; and second, it allows for situations in which preempting may not be rational. We can distinguish (at least) two types of situations in which S must rationally let EA's reasons for belief preempt his own:

(a) Cases of disagreement between S and EA.[40]
(b) Cases of agreement between S and EA, where S is entitled to suppose that his reasons for believing *p* are included among EA's reasons.

This analysis of PTW-B accommodates instances of (a), such as the doctor case and the RED CARDINAL DISAGREEMENT CASE, and instances of (b) such as the RED CARDINAL AGREEMENT CASE. However, condition (b) does not settle the RED CARDINAL SISTER AGREEMENT CASE, where I am entitled to suppose that my sister is not sensitive to birds' sounds, and hence that aggregating her reasons with mine is more rational than just preempting.

Someone might object to (a), arguing that not every case of disagreement requires S to preempt his own reasons: indeed, there might well be situations in which EA gets something wrong and S should have held on to his original belief. It is important to notice though that such circumstances would call EA's status of epistemic authority into question, rather than S's rational commitment to preempting. Indeed, when you disagree with an EA on something, if you have conscientiously judged that she is an authority—that she is more likely to get the truth than you are—it is still rational for you to believe that you are more likely to be wrong than she is, and hence to adopt her belief.[41] However, if you find out that she is wrong (and you are

[40] Notice that here I only consider classical cases of non-peer disagreement, such as Goldman's novice/expert problem (2001). For the sake of simplicity, I do not take into account more complex cases of disagreement between experts, such as Goldman's novice/2-expert problem, where S has to face the judgment of two authorities who disagree with each other. See Jäger (2016) and Zagzebski (2016) for further considerations on this particular case.
[41] In fact, situations of peer disagreement differ from cases of non-peer disagreement in an important respect: the former might give S a *prima facie* reason to believe that the interlocutor is not an epistemic peer, whereas in the latter S still has a *prima facie* reason to defer to EA, i.e., the mere fact that she is an epistemic superior.

correct), then you should start wondering whether it is still rational to consider her an epistemic authority, that is, whether adopting her belief maximizes your chances of getting the truth.[42]

Arnon Keren (2014b) offers an argument against the functioning of Zagzebski's Preemption Thesis, which challenges condition (a). He recalls that Zagzebski's justification of the rationality of preempting in cases of disagreement is grounded in Raz's argument for practical authority. This argument can be summarized as follows:

> Suppose I can identify a range of cases in which I am wrong more than the putative authority. Suppose I decide because of this to tilt the balance in all those cases in favor of its solution. That is, in every case I will first make up my own mind independently of the "authority's" verdict, and then, in those cases in which my judgment differs from its, I will add a certain weight to the solution favored by it, on the ground that it, the authority, knows better than I. This procedure will reverse my independent judgment in a certain proportion of the cases. Sometimes even after giving the argument favoured by the authority an extra weight it will not win. On other occasions the additional weight will make all the difference. How will I fare under this procedure? If, as we are assuming, there is no other relevant information available, then we can expect that in the cases in which I endorse the authority's judgment my rate of mistakes declines and equals that of the authority. In the cases in which even now I contradict the authority's judgment the rate of my mistakes remains unchanged, i.e., greater than that of the authority. This shows that only by allowing the authority's judgment to pre-empt mine altogether will I succeed in improving my performance and bringing it to the level of the authority (Raz 1988: 68-69).

Keren points out that the Razian argument is meant to explain preemption in cases of practical authority, where S has just two options—namely, doing an action *a* and not doing *a*. In contrast, in the epistemic case—relevant to Zagzebski—S has at least three options, namely "believing *p*", "believing not-*p*", and "suspending judgment", and two epistemic goals, namely "forming a true belief" and "avoiding a false belief". Given these elements, Keren argues that S's suspending judgment "in at least some proportion of the cases" of disagreement between S

[42] As both Jäger and Zagzebski illustrate, there are cases in which authorities get things wrong, change their minds about something, or reveal themselves to be less authoritative than laypeople expect. Therefore, there is no reason why we should deny some sort of fallibilism on S's part and grant the possibility of S's re-considering EA's status of epistemic authority.

and EA—instead of preempting—lowers S's probability of error not only below that of S's independent judgment, but also below that of EA (see 2014b: 74). If S has better scores than EA in avoiding falsehood by withholding judgment, then it is not the case that S will be more likely to form a true belief and avoid a false one if S lets EA's judgment preempt his own. Thus, we should conclude with Keren that Zagzebski's account of the rationality of preempting in cases of disagreement fails and hence that PTW-B should not apply to condition (a).

However, Keren's argument is susceptible to criticism. Zagzebski (2014) acknowledges the importance of considering "withholding judgment" an available strategy for S, yet she replies that Keren's objection loses its strength once we show that if S has three options, EA should have three options too. In other words, if we concede that EA can suggest that S suspend judgment then preempting still assures the best track record in satisfying both epistemic goals, for it is rational for S to withhold judgment whenever EA says to do so.

Taking a different line of argument, I want to show that, despite its initial plausibility, Keren's strategy leads to unwelcome and counterintuitive results. Let us assume that in cases of disagreement between S and EA, EA is more likely to be right than wrong.[43] If so, then we can easily demonstrate that by withholding judgment S is more likely to fail to acquire true beliefs than to avoid false ones. Thus, it is more likely that suspending judgment will lead S to *not* reach the former goal than to reach the latter.

Although Keren is merely concerned with proving that suspension of belief lowers S's probability of error, he should not underestimate the (potential) loss of true beliefs that withholding judgment involves. For, assuming that forming true beliefs and avoiding false ones are equally relevant epistemic goals, as appears to be the case in Zagzebski's view, S's overall score with respect to the two goals considered together is lower when he suspends judgment than it is when he preemptively adopts EA's belief. In other words, withholding judgment leads S to obtain worse overall scores than he does by preempting. These results strike me as good reasons to avoid suspension of belief, especially because both Keren and Zagzebski postulate that the epistemic goals are equally relevant. Therefore, despite the technical soundness of his response to Zagzebski, Keren's strategy still looks implausible due to his underestimation of the costs of withholding judgment.

Now, let us examine (ii). Despite the examples supporting Jäger's argument against the Preemption Thesis, i.e., RED CARDINAL AGREEMENT CASE and RED CARDINAL SISTER

[43] This assumption is in line with the very idea that EA is an epistemic authority for S. Conceding that EA "is much more likely to be mistaken" when she disagrees with S about *p* compared to cases on which they agree, as Keren does (2014b: 74), does not entail that EA should be more likely to be wrong than right as to whether *p* is the case. If that were the case, then it would seem unreasonable to consider EA as an authority in that domain.

AGREEMENT CASE, I think Jäger is wrong to claim that preempting is in itself incompatible with the authority of understanding. Consider again a version of the example I have discussed in §2.4 involving a grandmother and her grandson:

> GRANDMOTHER FISH CASE. Suppose that a child truly believes that fish breathe, but that his belief is grounded in some fantasy theory that they just need to surface and breathe once per day, and that they do it at night, when nobody can notice and catch them. Suppose also that his grandmother, despite lacking any particular competence in zoology, gives him a more reasonable explanation for how fish breathe and that she persuades him.[44]

In such a case, the child is in a position to disregard his initial understanding of the fish respiratory system and to adopt his grandmother's, for he realizes that he is more likely to gain a better understanding of this subject if he trusts his grandmother than if he sticks to his own judgment, as suggested by JAU1.

However, to demonstrate that it is rational for S to let EA's understanding preempt his own, I first need to resist the objection that "understanding can't simply be given to another in the way knowledge can" (Gordon 2017: 298), that is, via testimony. Philosophers like Emma C. Gordon, Allison Hills, and Linda Zagzebski have already defended what Kenneth Boyd recently called the *Indirectness Thesis*, according to which "[t]estimony cannot be a direct source of understanding: at best, it can be an indirect source of understanding by laying the groundwork for potential understanding" (Gordon 2017: 305).[45]

Following Boyd, we can conceive the notion of understanding as consisting of two main components: the *informational* component and the *grasping* component. The former amounts to possessing the apt information we need for understanding a subject matter *x*; hence it can easily be transmitted via testimony, as the teacher does when explaining the causes of the Iraq War to the student and as the grandmother does when she tells the child about how fish breathe. The latter component involves both the apt exercise of some intellectual ability on S's part and "the relationships between the thing understood and either the reasons that make that thing true or propositions that are related to it" (Boyd 2017: 114). I shall say more about the notion of grasping in §4.2.1: for the time being, suffice it to say that grasping encompasses the abilities

[44] We do not even demand that she knows much about fish respiratory systems; it is merely required that her explanation better fits other true beliefs related to mammals, gills, personal experiences, or narratives about fish life, etc.

[45] See also Hills (2009) and Zagzebski (2008: 145). For an argument in favor of the thesis that testimony can generate—rather than merely transmit—understanding, see Malfatti (2019b).

to follow an explanation of some issue, connect that explanation with the background information one possesses on the topic, assess the grounds of that body of information and what follows from it. As should be evident, testimony conveys information, not abilities: hence, it looks as though the speaker's testimony cannot provide the hearer with the abilities that are required for (satisfying the) grasping (component). This is why, as the Indirectness Thesis states, testimony can at best count as an indirect source of understanding: for the testimonial exchange to be successful, the hearer has to possess and make use of his grasping skills.

Nonetheless, Boyd claims that this thesis only applies to a specific kind of understanding, i.e., *difficult understanding*—that is, to cases requiring long-term effort and the exercise of many abilities, such as understanding the tiki-taka strategy in football or Navies-Stokes equations in fluid mechanics. Other kinds of understanding have more relaxed demands. Boyd considers two further kinds of understanding, which he calls *easy understanding* and *easy$_s$ understanding*. Anyone possessing a general background familiarity with the subject matter at stake can achieve easy understanding simply "by making sense of" the interlocutor's utterances, as the student and the child can do in the aforementioned examples (see Boyd 2017: 119). On the other hand, easy$_s$ understanding works as easy understanding for a particular range of subjects who have specific background familiarity with a more complex issue and who have developed the necessary abilities to grasp the information they receive. For example, understanding the Gettier problem is a difficult challenge for a young philosophy student, yet it is definitely easier for a philosophy professor.[46]

To demonstrate that one can achieve these kinds of understanding via testimony, Boyd argues that satisfying their grasping component does not demand too much effort from the subject, since making sense of the testimony itself suffices to allow one to grasp the relationships between bodies of information. Thus, even though the grasping component, technically speaking, does not transmit from one person to another it is still true that the mere fact that one conveys her understanding via testimony suffices to let the hearer acquire that understanding. For making sense of the speaker's utterances entails fulfilling both the informational and the grasping component at once. Therefore, we can conclude that, pace

[46] Interestingly enough, Boyd's distinction between easy and difficult understanding corresponds to Pritchard's account of weak and strong epistemic achievements (2010). However, since on Pritchard's view understanding constitutes a form of strong achievement, it might be objected that Boyd's notion of easy understanding fails to capture the core of what it takes to possess understanding. Discussing this objection in detail is beyond the scope of this chapter. I shall limit myself to noting that the thesis that preemption let novices acquire epistemic goods (understanding included) appears to be compatible with the possibility that their achievements are only of a *weak* form—namely, a success through ability that does not involve overcoming a significant obstacle or the exercise of a significant level of ability (Pritchard 2010: 70). It might be the case that for a novice to obtain a *strong* achievement through an exchange with an epistemically superior individual, it is necessary that the novice weighs reasons in favor and against the authority's testimony.

Gordon, easy understanding and $easy_s$ understanding can be given to another as knowledge can, that is, via testimony.

Furthermore, easy understanding and $easy_s$ understanding allow for preemption. Preemptive transmission of understanding between EA and S requires S to disregard whatever understanding of a subject matter *x* he possesses and adopt EA's understanding of *x* without considering her testimony as evidence to weigh against his own. Boyd's analysis explains why preempting understanding is possible, as acquiring easy understanding and $easy_s$ understanding by making sense of another's words does not require the cognitive effort and the exercise of intellectual abilities that weighing evidence demands.

Some might want to resist the idea that it is possible to preempt understanding by arguing that understanding does not satisfy Jäger's account of preemption in terms of *proper basing*.[47] According to Jäger, preemption demands that "the authority's belief should typically constitute a reason *for which* one adopts" her belief (2016: 175). In the case of understanding, holding that the authority's understanding of *x* constitutes a reason for which one adopts her understanding of *x* seems (at best) unclear. For adopting another's understanding always requires that I grasp the inner relationships between the various parts of her testimony as well as the relationship between the testimony itself and my background knowledge concerning *x*.

However, Boyd's distinction between the two components of understanding suggests a way to accommodate this worry and to argue that Jäger's proper basing account explains preemption in the case of understanding as well. We should not be surprised that proper basing does not apply to the grasping component, as we have already shown that this component cannot be transmitted from one person to another. In contrast, this account does apply to the informational component: the epistemic reason *for which* S accepts the information conveyed by the authority via testimony—and therefore satisfy the informational component—is exactly the fact that it is EA's testimony of her understanding. Thus, we can account for preempting understanding in terms of proper basing, as the communicable component of understanding satisfies Jäger's requirements.

Before moving to the analysis of the circumstances in which preempting EA's understanding is rational for S, I think it is worth mentioning one more point about the role of the authority of understanding. Letting another's understanding preempt my own is possible only if the interlocutor is able to convey her understanding (or some portion of her understanding) in such a way that I can grasp the relevant relationships between bodies of

[47] Jäger defines proper basing as follows: "S's (graded) belief B is properly based on a given ground (or set of grounds) G iff (i) G is the ground for which S holds B and (ii) S has a true and rational belief to the effect that G sufficiently supports B" (2016: 174).

information simply by making sense of her utterances—as both easy understanding and $easy_S$ understanding demand. Such a condition casts light on a specific ability that authorities of understanding need to display: namely, sensitivity to S's epistemic resources. I will say more about the authority's virtues in section 4.

Now it is possible to introduce a version of PTW and clarify how preempting works in cases of the authority of understanding as follows:

> *Preemption Thesis Weak for the Authority of Understanding (PTW-U).*
> There are circumstances in which the fact that an EA has some understanding of a subject matter *x* is a reason for me to accept her understanding of *x* that replaces my previous understanding of *x* and is not simply added to it.

The two main circumstances in which PTW-B justifies S's preemptively adopting EA's beliefs apply to PTW-U as well, thereby showing that a limited-preemptionist view can accommodate the authority of understanding. When S disagrees with EA about *x*, that is, condition (a), preempting may well be the most rational option for him, as in the GRANDMOTHER FISH CASE, where the grandmother clearly has a broader, deeper, and more significant understanding than the child. Moreover, we can imagine that the grandmother has developed the necessary sensitivity to her grandson's epistemic resources such that she is able to convey to him an easy understanding of how fish breathe.

Someone might object that there are cases of disagreement in which preempting is not rational, since EA is wrong and S is right. As in PTB-W, similar cases threaten EA's status of authority, not the rationality of preempting. Preempting understanding is rational insofar as EA's understanding of *x* fares better than S's understanding of *x* on at least one of the three dimensions of understanding—i.e., breadth, depth, and significance—without faring worse on any of them.

This rule applies to condition (b) as well: when EA's understanding of some subject matter *x* is just broader, deeper, or more significant than S's, and hence the informational component of EA's understanding of *x* already includes all the information about *x* that S possesses, it is rational for S to preempt. This is what happens in the HISTORY PROFESSOR CASE, where the teacher's broader understanding of the Iraq War includes the student's explanation for the US military intervention in Iraq but places it within a more comprehensive story of the causes of conflict. Therefore, both JAU1 and PTW-U apply to the HISTORY PROFESSOR CASE, for it is

rational for the student to disregard his previous understanding and preemptively adopt the teacher's.

However, as condition (b) of PTW-B is not in a position to account for the RED CARDINAL SISTER AGREEMENT CASE, there are also cases where S's understanding is not completely included within EA's. Consider the following example:

> PHILOSOPHY PROFESSOR CASE. Suppose that I acquire, through intensive study, some understanding *U* of the traditional dichotomy between reductionism and non-reductionism in the epistemology of testimony (*x*). Suppose also that my professor has a broader understanding of *x* (U^I), as she has an explanation for how recent hybrid theories intend to overcome both traditional views, yet not a deeper and more significant one, since my research on reductionism and non-reductionism has been extremely thorough.

In situations like the PHILOSOPHY PROFESSOR CASE, even if it were possible to preempt, it would still not be rational. Indeed, since acquiring U^I via preemption entails that my own understanding of *x* (*U*) "stops operating" (Jäger 2016: 175), by doing so I would be acquiring an understanding of the relationships between the traditional views and hybrid theories from the professor without grounding this new information on the more detailed story of *x* that I already possess. Suppose JAU1 still applies to the PHILOSOPHY PROFESSOR CASE insofar as I acknowledge that the professor has a broader understanding U^I than I have and I judge that it is more likely that I gain a better understanding of *x* if I believe what the professor believes. In this case, it would still be wise not to preempt, because basing the new information about hybrid theories of testimony provided by the professor on my deeper and more significant understanding of *x* would allow me to establish more robust explanatory relationships between traditional views and hybrid theories than the ones I would acquire if I let the professor's understanding preempt mine. Therefore, PTW-U does not apply to the PHILOSOPHY PROFESSOR CASE.

To summarize, Claim 2 allowed me to defend a weak form of the Preemption Thesis, that is, to conditionalize the application of this thesis upon specific circumstances in which preempting is still the most rational thing to do for S, both in cases of the authority of belief and in cases of the authority of understanding.

3.4.3 The Limits of Jäger's Account

A final point against Jäger's account of epistemic authority can be found in Claim 3:

> *Claim 3*: Socratic authority (SA) should represent a special case of authority, rather than an alternative account of authority.

The argument in support of Claim 3 is quite straightforward. What Jäger has in mind when he offers his alternative account of epistemic authority is Socrates' maieutic ability. This ability is such that Socrates does not tell anyone that such and such is the case: he merely asks questions that guide interlocutors to understand things for themselves. Hence, Socrates' questions provide people with reasons to expand their understanding in a very *indirect* way. On the one hand, Socrates' method rightly places him and those who adopt the maieutic method among the authorities of understanding. But on the other hand, his higher ratio of true to false beliefs than the interlocutor *plus* his ability to recognize false beliefs within the corpus of the interlocutor's beliefs should make him an authority of belief as well.[48] The fact that SA has this twofold authoritative role for the layperson justifies considering it a special case of authority. Indeed, Jäger's view illustrates a specific, *viz.* indirect, way in which someone can fulfill what I take to be the general scope of epistemic authority, that is, responding to S's various epistemic needs.

Nevertheless, it would be a mistake to define epistemic authority merely in terms of Socratic authority, for two main reasons: first, Jäger's account rules out authorities of understanding that lack maieutic ability and therefore guide laypeople in a more direct way, e.g., the grandmother who straightforwardly presents the child with a better explanatory story of how fish breathe, rather than leading him to work out that his understanding is wrong through critical questions. Second, it is insensitive to any distinction between authorities of belief and mere experts, neither of which is considered an epistemic authority on his view. Thus, Jäger's account severely restricts the opportunities for a layperson to defer to an epistemic authority.

Some might believe that this consequence of Jäger's view is unproblematic for the agenda of social epistemology. For even if experts and authorities who lack maieutic ability do not fulfill his account of authority, we can still acquire testimonial knowledge from them. Notice, though, that this reply is flawed. Many epistemologists concede that someone's testimony provides the hearer with weighing reasons. Furthermore, Jäger holds that this view may well apply to the interaction between the authority and the layperson. However, the weighing-model

[48] See section 3.5 for further clarifications on Socratic authority.

of epistemic authority differs from testimonial exchanges among epistemic peers in an important respect. According to this model, it is precisely the fact that I acknowledge my interlocutor's epistemic authority that explains why EA's testimony that *p* provides me with a stronger reason to believe that *p* than the one I would acquire from an epistemic peer's testimony that *p*.

Thus, if we follow Jäger and deny experts and authorities who lack maieutic ability the status of epistemic authorities, it is doubtful whether an expert's testimony that *p* provides us with the strong reason to believe *p* that we expect to possess whenever we deal with such epistemically superior subjects. This is bad news for social epistemologists aiming at explaining the value of our dependence on epistemic authorities, because Jäger's view fails to appreciate an important epistemic advantage to deferring to authorities, rather than to epistemic peers. Thus, unless Jäger offers us better reasons for endorsing this restrictive view, we can conclude that Socratic authority cannot represent a plausible account of epistemic authority; rather, it constitutes a special kind of authority.

We can draw further evidence for this conclusion from considerations about the Preemption Thesis. Jäger does not account for any authority whose exercise entails issuing preemptive reasons to the interlocutor. But if the reasons I offered in support of PTW-B and PTW-U are compelling, then Jäger owes us an argument for rejecting the idea that in specific circumstances epistemic authorities can provide the interlocutor with preemptive reasons. Furthermore, no version of the Preemption Thesis can work for Socratic authorities. To understand why, consider the following case, involving the grandmother and her grandson again:

> GRANDMOTHER BOAT CASE. Suppose the inquiring child notices a huge yacht moored in the harbor and forms the incorrect belief that boats float because fish hold them up once they see their shadow. This time, the grandmother wants to let the child correct his opinion by himself. Hence, she asks him how many fish are holding up the yacht and he replies that there must be hundreds of medium-size fish below it. Then the grandmother asks: "But why are there no fishermen trying to catch them all?" Now, her grandson is confused and unsure of what to believe. So to help him, she asks whether fish also hold him up when he lies down in the water. He rightly replies: "No, I don't see fish around me when I am swimming. It's just the water that makes me float!" Thus, the child now believes that his original explanation was wrong and that a better explanation does not involve fish holding up objects to make them float in water.

This example shows how the interaction between a SA and a layperson might work. Indeed, the grandmother does not use her authority to present the child with a better understanding; rather she challenges him with critical questions so as to lead him to reassess his own judgment. It seems evident that Socratic authority leaves no room for preempting: if SA cares about a layperson's intellectual development and limits herself to putting S in a better position to improve his understanding by himself, then she does not give any true and clear-cut testimony of her understanding that S might preemptively adopt. Therefore, we can conclude that SA should be merely conceived as a special subset of epistemic authority including both the authority of belief and the authority of understanding.

As a final remark, notice that there may be cases of Socratic authority that violate Zagzebski's assumption that an epistemically conscientious S needs to acknowledge the interlocutor's authority over him. In other words, it is not necessary in the case of Socratic authority that S possesses a reason to believe that SA is more likely to have a true belief or a better understanding than S is. This violation is legitimate because the maieutic method does not require that S defer to SA, neither by preempting nor by weighing the reasons provided via testimony. Rather, S enhances his understanding simply by reflecting on the interlocutor's questions.

The discussion of the three claims introduced in this section has set the ground for the broader theory of epistemic authority I will develop in section 3.6. Claim 1 provided reasons for broadening Zagzebski's account so as to elicit S's appeal to an authority to expand his understanding of some subject matter. Claim 2 explained why it is wise to maintain a weak version of the Preemption Thesis (i.e., PTW-B) and to admit that it also applies to the authority of understanding (i.e., PTW-U). Finally, Claim 3 made clear that a comprehensive account of epistemic authority ought to define Jäger's Socratic authority merely as a special case of authority of both belief and understanding.

3.5 Lackey against Preempting and The Authority-as-Advisor Model

Before putting together the views of an expert and an epistemic authority that I have defended so far, I shall consider one more set of objections to the Preemption Thesis, namely one that has been recently put forth by Jennifer Lackey (2018). Discussing Lackey's view is relevant

not only to shed light on the pros and cons of the preemptionist view of authority, but also to take into consideration an alternative model of epistemic authority, according to which the function of (experts and) authorities is that of advising laypeople by offering them the evidence they might lack in a given domain.

3.5.1 The Track Record Argument

Let us begin by introducing and discussing in more detail the Raz-Zagzebski *track record argument* that constitutes the main target of Lackey's critique. As anticipated in §3.4.2, Raz has it that once S acknowledges that he is more likely to be wrong than the authority in some domain, endorsing the authority's judgment makes S's rate of mistakes decline and equal that of the authority. In contrast, when S disagrees with the authority and sticks to his own judgment, his rate of mistakes remains greater than that of the authority. Hence, the most rational strategy for S to improve their performance and get as close as possible to the authority's performance is to let the authority's judgment preempt S's.

Zagzebski further argues in favor of this preemptionist argument by attempting to accommodate two facts that make preempting particularly difficult to accept and adopt by human beings. The first fact has to do with our psychological resistance to accepting that, in some circumstances, we are better off preempting than deploying any other decision-making strategy—e.g., weighing reasons in favor of and against a given course of action.

Zagzebski draws our attention to empirical studies conducted by Mlodinow (2008), according to which the human tendency to proportionate their decisions based on the probability that an event occurs makes us do the right thing (or get to the truth) less often than we could. Mlodinow's experiments in fact compare human behavior with the conduct of rats and pigeons: as it turns out, rats are better "preemptioners" than human beings and, for this reason, rats end up better off than we do. More specifically, if an event *x* has a 0.7 probability of occurring, we will decide to undertake the actions that allow *x* to occur only 70 percent of the time. Conversely, when rats understand that a given course of action leads them to an objective the majority of the time—e.g., predicting that a green, rather than a red, light will flash—they will always stick to that course of action. The conduct of these animals ensures that they get their reward 70 percent of the time, whereas the proportioning attitude of human beings leads us to worse results. For besides the fact that some of the times when we stick to the most promising course of action—e.g., predicting that the green light will flash—our choice will not pay off as

it happens to rats and pigeons, the human tendency to match probabilities also leads us to deviate from the most promising course of action when we should have stuck to it—e.g., predicting that the red light will flash when, in fact, the green one does.

The second source of resistance toward preempting pertains to the fact that this attitude, in the context of epistemic authority, pairs with what H. L. A. Hart calls *content-independence* (1990: 101), according to which "[a]n authoritative utterance gives the subject a reason to follow the directive which is such that there is no direct connection between the reason and the action for which it is a reason" (Zagzebski 2012: 106). In other words, part of the reason why human beings struggle to accept preempting is that its rationality does not seem to depend on how plausible or reasonable the testimony of the authority looks to the hearer.

To explain away this alleged problem with preempting, Zagzebski invites us to consider the following example:

> Suppose your physician tells you to take 4,000 pills an hour for the rest of your life. I assume that you trust your belief that you should not take so many pills more than you trust your judgment that your physician is an authoritative guide to your health. To determine whether the physician is a better guide to your health than you are, you have a right to take into consideration anything that you find trustworthy when you are conscientious. But as long as you conscientiously think the physician is a better guide, you have reason to take the physician's directive as one that preempts your own decision about what you should do in that domain. And, of course, the same point applies to your *belief* that you ought to take 4,000 pills an hour (2012: 116).

The moral of this example, as Zagzebski herself admits, is that it might be the case that the outrageousness of an epistemic authority's testimony can make us doubt her epistemic status, but it should not challenge the soundness of the Preemption Thesis. The alleged problem with preempting hides a more substantial problem with the Hartian content-independence principle. Basically, we fail to comply with the Preemption Thesis because we tend to trust epistemic authorities only when their testimony survives a process of weighing reasons in favor of and against what they tell us: in other words, we cannot separate the judgment about the epistemic superiority of our interlocutor from the judgment about the plausibility of what she communicates to us.

3.5.2 Lackey's Objections

In this section, I shall discuss two specific objections that Lackey raises against the track record argument, which I respectively call *the problem of outrageous testimony* and *the problem of the authority's mistakes*.

The problem of outrageous testimony has to do with the 4,000 pills example. According to Lackey, cases of an authority's outrageous testimony reveal that the track record argument is flawed because it seems evident that we should not follow the authority's advice in such circumstances. Furthermore, Zagzebski cannot appeal to the fact that in cases like the 4,000 pills example we might have evidence that our interlocutor has lost the epistemic authority she had on some matter. For conceding this simply amounts to doing the opposite of preempting, namely *weighing* reasons, reflecting on the testimony we have received based on our background evidence. Thus, according to Lackey (2018: 235-236), the Preemption Thesis turns out to be incompatible with the track record argument.

The problem of the authority's mistakes has to do with determining how to avoid the layperson's track record getting worse. According to Lackey, more rational policies than preempting would mandate to follow the advice of an authority except when: one is certain that the authority is wrong; one knows that the authority is wrong; what the authority says is highly doubtful (238). Adopting any of these policies, on Lackey's view, would ensure that our track record would be better than not only the track record of those who preempt, but also the track record of rats and pigeons. Moreover, we would also be overall better off for we would not need to set aside our evidential background.

In the remainder of this section, I will try to show why these objections to the track record argument do not undermine the preemptionist view of epistemic authority. Let us consider the problem of outrageous testimony first. Lackey argues that S cannot appeal to his evidence to raise doubts about the epistemic authority of a speaker who has just offered an outrageous piece of testimony because doing this requires weighing the reasons in favor of and against accepting the testimony. However, it is not clear that this conclusion follows from the premises. The preemptionist could well reply that acknowledging the outrageousness of EA's testimony merely requires that S be an epistemically conscientious subject, that is, someone who is sensitive to potential defeaters for the testimony he receives.

Let us consider the 4,000 pills example again. A conscientious patient need not make any heavy weighing to realize that the number of pills he is being asked to take is absurd: 4,000

pills per hour amounts to more than 66 pills per minute, thus more than a pill per second. Nor should it be hard for him to recall that he has never been prescribed anything remotely close to what the doctor is telling him right now. Whether or not the patient does the math or makes the inference about his past experience, what matters is that his sensitivity to defeaters suffices to make him realize that the physician's testimony cannot be accepted without further investigation, thereby preventing him from adopting a preemptive attitude.

A similar line of defense has been recently pursued by Constantin and Grundmann (2018), according to which the pills case is one in which the outrageousness of the testimony provides S with a defeater, that is, an indicator "*lying outside the domain* of medicine that there is something wrong with the recommendation and that it is therefore not preempted" (17). It is thus the existence of a domain-independent reason against trusting the authority that allows the hearer not to preempt.

In sum, the preemptionist reply to the problem of outrageous testimony is that conscientiously relying on epistemic authority neither requires that S always weighs the reasons offered by the authority's testimony, as Lackey would commend, nor that S blindly takes for granted whatever the authority recommends. Rather, a plausible preemptionist view suggests that a conscientious subject be sensitive to defeaters, that is, willing and able to assess the reasons in favor of and against accepting an authority's testimony when the recommendation is blatantly outrageous. More specifically, here the expression "blatantly outrageous" identifies those cases in which the outrageousness of an authority's testimony is not generated by the content of her esoteric statements—that is, by statements concerning her domain of epistemic authority that are largely inaccessible to the lay interlocutor—but rather by the content or the implications of statements involving something outside the domain of expertise, on which even a lay hearer might have a say—for example, a rather practical issue like the number of pills one could be reasonably expected to take in one day.

The case of esoteric statements within EA's domain of expertise is one in which S's sensitivity to defeaters cannot activate because the content of EA's testimony goes beyond S's epistemic competence. In a case like this, to the extent that S has reason to consider EA epistemically superior to themselves, it is rational for S to rely on EA and preempt, despite the fact that doing so exposes S to the risk of being fooled or misled by an insincere or unreliable interlocutor. What makes this attitude rational is the fact that S trusts his epistemic faculties which make him judge that EA possesses epistemic authority in a given domain.

An example proposed by Constantin and Grundmann (2018: 17) helps to clarify this point. They ask us to consider the case of physicists, who make outrageous claims all the time from

the perspective of laypeople. For example, experts in quantum mechanics contend that objects do not possess a definite location and may materialize spontaneously: no doubt these claims may sound crazy to laypeople, but we do not take their outrageousness to provide us with reason against believing what physics teaches us about the world. Quite to the contrary, we trust physicists to find out more and more about the behavior of matter and light on behalf of the entire epistemic community, and we are rational in believing what they tell us—most of which constitutes, for the majority of people, esoteric statements.

I have addressed Lackey's first objection by illustrating that there are cases in which the outrageousness of an authority's testimony authorizes S to refrain from preempting without committing us to abandon preemptionism as a theory of epistemic authority. Let us now consider the problem of the authority's mistakes.

Here the best line of defense of preemptionism is to acknowledge that there seem to be no reasons why a preemptionist should resist Lackey's constrained version of the track record argument. Cases in which S knows or is certain that the authority is wrong or in which what EA says is highly doubtful constitute situations in which EA is not as epistemically conscientious as she is supposed to be. The track record argument ought not to mandate that we follow EA's recommendation when we know it to be wrong. As a matter of fact, relying on authority is just a means to a more important end, namely getting to the truth, and believing EA's words whenever S knows her to be wrong would just lead S astray from his epistemic goals. In such circumstances, S has a conclusive reason to refrain from preempting and, possibly, to question EA's status of an epistemic authority.

I want to close this section with a more general remark about preemptionism in light of the objections I have just discussed. Lackey's critique is meant to warn us about the risks of becoming servile and gullible epistemic subjects who stop exercising our critical attitudes and blindly defer to epistemic authorities. For one thing, no epistemic good can arise from setting aside our epistemic abilities, which are the most powerful tool at our disposal to achieve our epistemic goals. For another, we also want to prevent alleged or fake epistemic authorities—e.g., gurus or charismatic leaders of terroristic and extremist groups—from having the chance to manipulate a large number of followers into believing what they want.

But as I have tried to show here, it is not the preemptionist view that we should be concerned with. Preemptionism does not require that we stop doing our epistemic work and use our abilities to the best of our capacities. Quite to the contrary, epistemically conscientious subjects should deploy their faculties and intellectual virtues to work out the best strategy to achieve their epistemic goals. This might well include letting an epistemic authority's reason

preempt one's own, if one is right to judge that the authority is more conscientious than one is in a given domain. But surely that does not come out of a refusal to be reflective and critical; rather, it arises out of a rational strategy to maximize epistemic achievements.

3.5.3 The Authority-as-Advisor Model

In her recent contribution to the debate about epistemic authority, Lackey also outlines an alternative account of authority that, in her view, should replace the preemptionist model. She calls this account the "expert-as-advisor" model (2018: 238) but, given what I have argued so far about the differences between experts and epistemic authorities, I shall stick to my labels and rename her view the "authority-as-advisor" model.

The important aspect of this model of epistemic authority is that it replaces the function of authorities on Zagzebski's account—namely the idea that EA provides S with true beliefs that S lacks—with another function, that is, the idea that an advisor "provides evidence for believing a given proposition and, in this way, offers guidance" (238). Paradigmatic examples of this model of epistemic authority involve expert testimonies in trials and ethics consultants in hospitals. The reason why we consider an expert testifying to court as an epistemic authority is not that she provides jurors with preemptive reasons for accepting what she reports, but rather that she puts jurors in a position to assess themselves whether to form a belief based on the evidence that she makes available to them. Similarly, in the case of the ethics consultants, it is not as though they have a normative power to set aside the beliefs and the epistemic reasons of those who speak to them and provide preemptive considerations. Rather, ethics consultants provide the hospital staff, patients, and their members with reasons that allow them to better understand the moral and human complexities of a particular situation.

Thus, on Lackey's authority-as-advisor model, the advisor's epistemic superiority goes well beyond a mere concern with their epistemic reliability. The advisor is not just someone who relieves us from the burden of seeking out conclusive reasons in favor of a given belief; they also enhance "our understanding of the matter" and exemplify important intellectual virtues (239). In particular, it looks as though the abilities "to listen attentively and receptively to the concerns and values of those around her, and to answer questions in a thoughtful and constructive way" (239) are a fundamental component of what it takes to be an epistemic authority on the authority-as-advisor model.

As I have noted in the case of Jäger's view, Lackey's account also makes an important contribution towards broadening a theory of epistemic authority and, in particular, a conception of the service that authorities can perform within an epistemic community. Their views highlight the weaknesses of Zagzebski's original account but will allow me to further stress the advantages of the pluralistic view of epistemic authority that I shall set out in the next section. Before getting to that, though, I shall briefly consider the problem of blatantly false testimony uttered by an epistemic authority and attempt to undermine Lackey's idea that the advisor-as-authority model handles this problem in a way that the preemptionist view cannot (241).

According to Lackey, on the advisor-as-authority model S is always in a position to reject EA's false testimony when its content clashes with S's evidential background, since an authority's testimony never screens off S's epistemic reasons. Instead, as I have explained in the last section while discussing the problem of outrageous testimony, Lackey thinks that the preemptionist view cannot give S this option, because his reasons are preempted by EA's testimony. Setting aside what I have offered in response to the problem of outrageous testimony, here I want to focus on a different problem that her view seems to incur. The mere fact that advisors do not prevent their interlocutors from weighing reasons does not guarantee that the novices are therefore able to reject a false (or outrageous) testimony. If that were true, those who are raised within closed-minded terroristic or extremist groups—more generally, within what Thi Nguyen calls *echo chambers* (2018)—guided by an epistemic subject who satisfies the requirements of the advisor-as-authority model should nonetheless figure out by themselves that the balance of reasons undermines their terroristic or extremist ideals.

But this is far from being so. Quite to the contrary, the proliferation of extremist groups and echo chambers (at least partly) depends on the abilities of their leaders to ensure that their followers lack the possibility to make the balance of reasons on their own. In fact, these leaders provide group members with false testimony and distort available evidence or screen off evidence that goes against their agenda. It should come as no surprise that even some disagreement-reinforcing mechanisms typical of echo chambers are implemented and sustained by the abilities of such leaders. Disagreement-reinforcing mechanisms go beyond a mere gatekeeping function, in that they not only prevent as much counterevidence and conflicting information as possible from being available within the group, but they also alert group members that counterevidence and false information is what they should expect from epistemically corrupt outsiders. Similar manipulation of information and evidence is just to be expected from leaders of extremely closed-minded communities that exploit the epistemic weaknesses of their members to pursue a specific agenda.

The take-home message of these considerations is that we should expect leaders of these particular kinds of epistemic communities to display several intellectual abilities that Lackey takes to be necessary for her advisor-as-authority model, yet to deploy them to indoctrinate their followers. Furthermore, it seems plausible to expect that such indoctrinating strategies affect the group members' ability to reject false or unjustified testimony. If this is so, it looks as though the problem of false (and outrageous) testimony does not depend so much on the specific normative power that we grant to epistemic authorities—be it the power of advising or that of providing preemptive reasons—but rather on how those who possess the required abilities to serve as epistemic authorities make use of their virtues.

On Lackey's view, the advisor-as-authority model has another advantage over preemptionist views concerning how the model handles cases of *isolated second-hand knowledge*, namely circumstances in which someone knows that *p* solely based on a single testimony that *p*—thus, second-hand—and knows nothing (or very little) relevant about the matter other than that *p*—which makes their knowledge isolated (Lackey 2011: 254). Expert assertions constitute a paradigmatic case against the thesis that isolated second-hand knowledge can be a sufficient epistemic basis for epistemically proper assertion.

Lackey considers a case in which an oncologist requires an ultrasound and MRI for a patient and, on her day off, a colleague tells her that the results of the test diagnose the patient with pancreatic cancer. Based on the sole testimony of the person in charge of the exams—that is, in absence of any report detailing the results of the tests—the oncologist tells her patient at their appointment that he has pancreatic cancer (253). According to Lackey, the example shows that experts need to have the capacity to advise laypeople because it is not epistemically permissible that they offer a piece of testimony without being in a position to provide any further explanation, that is, to address questions regarding their testimony. In the medical example, the oncologist is not justified in making her assertion because she cannot back up what she says with appropriate evidence: all she could refer to is her colleague's testimony, which does not suffice to put her in a position to answer the legitimate questions her patient might raise.

Thus, the example speaks against preemptionist views of epistemic authority, in that these views are committed to considering the oncologist's bold testimony as perfectly legitimate. According to Lackey, "[i]t would be perfectly acceptable on this view to not have anything to offer to a challenge or to simply respond, 'Because I said so' or 'Because I'm an authority on the matter.' After all, that the testimony is from an authority just *is* the reason for believing on this view" (2018: 241). Thus, this problem highlights a further advantage of the advisor-as-

authority model over preemptionism, in that on the former—unlike the latter—expert testimony from isolated second-hand knowledge does not constitute a legitimate way in which epistemic authorities can serve the epistemic community.

In the remainder of this section, I shall briefly argue why I think that preemptionism can be (at least partly) vindicated from Lackey's attack. I will postpone to the following section further considerations about how my proposed theory of epistemic authority handles Lackey's objection. Understanding how preemptionists might reply to this objection requires shedding light on a key conceptual difference between what the view mandates regarding the normative attitudes we should have towards epistemic authorities and what Lackey's argument suggests (or implies).

Preemptionism has it that it is rational to let an authority's testimony that *p* preempt the reasons we might have in favor of or against the truth of *p.* In other words, the view explains how expert testimony works and what the most rational response on the part of the novices is. However, *contra* Lackey, preemptionism does not suggest (or imply) that it would be perfectly acceptable for an authority to offer a simple "Because I'm an authority on the matter" as an explanation of her testimony. The fact that an authority's testimony provides the hearer with preemptive reasons does not entail that authorities are therefore free from the epistemic duty to (be able to) give proper explanations of what they say.

More specifically, the crux of the matter here is neither the isolation nor the second-hand nature of the knowledge that is passed along, but rather the fact that preemptionism, according to Lackey's analysis, completely overlooks an authority's epistemic duty to have reasons to offer in support of their testimony. On closer inspection, though, this diagnosis appears to rest on a misunderstanding. Clearly, a one-time episode in which an authority fails to address questions concerning her testimony does not undermine the preemptive nature of the reasons offered by her testimony. Nonetheless, it is plausible to contend that, in the long run, a conscientious novice will figure out that there is something wrong with an authority that always demands to be trusted without being willing or capable of providing any explanation for what she says. For although preemptionism has it that an authority's testimony provides novices with all they need to believe what the authority says, it does not require that they be completely blind towards clues of epistemically inappropriate conduct.

Quite to the contrary, as I have clarified at the beginning of this chapter, Zagzebski is clear that her whole theory involves two (or more) epistemically unequal-but-*conscientious* agents. Thus, it seems fair to expect that, despite having a limited competence in a domain, conscientious laypeople—that is, individuals who make use of their epistemic faculties as best

as they can—are sensitive to clues of epistemically inappropriate conduct on the part of their epistemic authorities. For one thing, this is mainly in their own epistemic—if not also practical—interest. For another, it does not require too much of them, as it has nothing to do with esoteric statements; rather, it concerns the more general practice of giving and receiving reasons in which all human beings are involved. In a similar way to other kinds of sensitivity, we can expect that this skill too works under the surface of conscious awareness and comes back to the surface (typically by instilling doubt in the novice's head) only when the interaction with the authority presents some odd feature. For these reasons, it looks as though the problem of expert testimony based on isolated second-hand knowledge fails to provide the advisor-as-authority model with any advantage over preemptionism.

The scope of these considerations was to set out a plausible line of reply to Lackey's objection on behalf of preemptionists. No matter how compelling it sounds, let me stress again that my theory of epistemic authority handles the objection in a different way, which I shall outline in the next section, right after introducing and explaining the basic tenets of my pluralistic view.

3.6 A Comprehensive Framework of Epistemic Authority

The arguments of the previous sections illustrated that the debate on epistemic authority calls for a more articulated theory than the ones offered by Zagzebski, Jäger, and Lackey. The main goal of this section is to set out and defend an alternative view of epistemic authority that adequately takes into account three dimensions: the different *functions* that epistemically superior agents may perform for the benefit of other members in their community; the different *abilities* that allow them to fulfill the respective services; and the different *epistemic goods* or *ends* that they can promote.

The most straightforward dimension to address is the third one: as I have argued earlier on in this chapter, epistemic authorities can help the novices acquire true beliefs but also understanding in a given domain and this is reflected in the aforementioned distinction between authorities of belief and authorities of understanding.

As regards the functions that epistemically superior agents may perform, the analysis offered so far in the dissertation has illustrated that there are at least two main services that they can fulfill, which I have respectively labeled *research-oriented function* and *novice-oriented function*. As argued in the previous chapter, the class of experts is the one that is supposed to

provide the former service, in that we expect that they contribute to the epistemic progress of the epistemic community by answering extant questions within their field of expertise and being on the lookout for new questions that might arise. On the other hand, we expect that epistemic authorities—more specifically, authorities of belief and authorities of understanding—serve the epistemic community by fulfilling the novice-oriented function, that is, by helping novices acquire new true beliefs or improve their understanding in some domain.

As anticipated in the previous chapter, the two main functions of epistemic authority require different sets of epistemic virtues. I have suggested distinguishing between:

> *Research-oriented abilities*: virtues that allow an expert or authority to exploit their fund of knowledge to find and face new problems in their field of expertise (e.g., thoroughness, intellectual perseverance, intellectual courage, self-scrutiny, intellectual creativity, open-mindedness, intellectual curiosity, autonomy, etc.); and
>
> *Novice-oriented abilities*: virtues that allow an expert or authority to properly address a layperson's epistemic dependency on them (e.g., sensitivity to S's needs, intellectual generosity, intellectual empathy, sensitivity to S's epistemic resources … maieutic ability).

Assuming this classification, I propose the following criterion for discriminating among different kinds of authority:

> *Classification Criterion for Epistemic Authority (CCEA)*. For one to be an expert, one is required to display research-oriented abilities; whereas, for one to be an epistemic authority (whether an authority of belief or an authority of understanding), one is required to display novice-oriented abilities.

Both the ability requirements and the criterion I am endorsing here need some clarification. First, the list of novice-oriented abilities goes from the easiest virtue to display, i.e., sensitivity to S's needs, to the hardest virtue to display, i.e., maieutic ability. This order of the novice-oriented abilities casts more light on the limits of Jäger's view, as he is committed to denying that those who lack maieutic ability can be epistemic authorities, despite their displaying (or not displaying) other novice-oriented abilities.

Second, sensitivity to S's needs plays a fundamental role within this framework because it allows EA to acknowledge S's dependency on her and to properly address his specific demand in a given situation. Someone like Shyler, who lacks this novice-oriented ability that operates only within a relationship between two (or more) subjects, cannot be an EA for the interlocutor. However, CCEA allows us to consider her an expert insofar as she possesses a sufficient degree of research-oriented abilities, which are not necessary features of epistemic authorities.[49]

Third, in order to determine whether EA counts as an authority of understanding for S rather than as a mere authority of belief, we should ask whether or not the sensitivity to S's needs activates further novice-oriented abilities possessed by EA. Indeed, it is plausible to suppose that an authority of understanding, who aims at providing better service to S than merely providing him with true beliefs, needs to display a broad range of intellectual virtues. In other words, the more someone is willing to help S to settle the matter by himself, the more she moves toward the authority of understanding set on the spectrum of epistemic authority. This is what happens, for example, if we move from the red-cardinal cases involving the ornithologist and my sister to the grandmother cases. The former are situations in which the epistemic authority merely has to grasp the epistemic need of the interlocutor and be willing to solve his problem by offering him a true belief that he lacks. The latter are situations in which further virtues are necessary for the epistemic authority to provide the service that the interlocutor needs. As I have argued in §3.4, besides caring for the epistemic wellbeing of the kid, the grandmother has to empathize with his epistemic situation, make use of her epistemic resources to ensure that her advice matches his epistemic possibilities, thereby displaying both intellectual generosity and maieutic ability.

Furthermore, it looks as though an epistemic authority's concern for the epistemic welfare of her interlocutor depends on and grows proportionally to the breadth and depth of the set of novice-oriented abilities that she displays. As Duncan Pritchard clarifies in the passage below, it looks as though fostering understanding and promoting the development of intellectual virtues are strongly related:

> The development of understanding and the development of intellectual virtue go hand-in-hand and are mutually reinforcing: Enhancing the intellectual virtues better positions one to gain understanding (so to ask the right questions, to seek out the

[49] Notice here that I am still pointing to the subject-independent dimension of expertise (see section 1). In order for someone to have more expertise than I do in some domain (i.e., subject-dependent notion of expert), it is not necessary that they display a high level of research-oriented abilities. All that matters is that they have more accurate information than I have in that domain, as held by Coady (2012).

answers, and so on), and in gaining understanding one will thereby be pursuing strategies, which will likely enhance one's intellectual virtues (2016: 117-118).

My theory of epistemic authority captures this feature of the relationship between understanding and being intellectually virtuous, in that it accommodates the idea that helping novices improve their understanding of some domain on their own is the best service an epistemic authority can provide, namely one that goes beyond merely providing true beliefs or understanding that the novice lacks. However, an important advantage of my theory is that it does so as Jäger and Lackey have suggested while accounting for the fact that, in some circumstances, handing on information that the novice should accept via preemption is the only thing that an epistemic authority can offer and hence what we expect them to be doing.

As a final remark, notice that since the aforementioned categories concern virtues, that is, dispositions that a subject acquires and develops by degrees, it should not be surprising that CCEA can only offer thick conditions that work as cues for distinguishing among different kinds of epistemic authority rather than a direct discriminating formula.

CCEA and the clarificatory remarks just offered allow us to refine the three central notions of my theory of epistemic authority as follows:

Expert. A subject A is an expert in domain D (for a subject S) iff:

(1) A has better understanding than the majority of people do in D;

(2) A possesses *research-oriented abilities*.

Authority of belief (AofB). A subject A is an AofB in domain D for a subject S iff:

(1) A is more conscientious than S—who considers her to be an EA—in D;

(2) There is an ongoing relationship between A and S;

(3) A possesses and makes use of *sensitivity to S's needs*.

Authority of understanding (AofU). A subject A is an AofU in domain D for a subject S iff:

(1) A is more conscientious than S—who considers her to be an EA—in D;

(2) There is an ongoing relationship between A and S;

(3) A possesses and makes extensive use of *novice-oriented abilities*.

The conscientiousness-requirement of authorities of belief and understanding calls for further clarification. Those who still worry that the notion of epistemic conscientiousness remains relatively obscure can replace condition (1) with a condition stating that (1*) A is better epistemically positioned than S—who considers her to be an EA—in D.

The notion of *epistemic position* constitutes the weakest marker of superiority in the epistemic domain. It has to do with an epistemic agent's internal conditions—i.e., their psychological and physical situation—as well as with external conditions—pertaining to, e.g., environmental and contextual features of the testimonial exchange—that may affect one's opportunity to access relevant information.[50] According to this condition, it might be the case that a subject A is better positioned than a subject S in a given domain not because A is, in general, more reliable or epistemically virtuous than S, but because, for example, S is under the effect of some powerful drug that impairs his epistemic faculties; S displays a bias that obstructs his ability to properly assess the evidence in the domain; or A finds herself to be in the most appropriate spatial-temporal position to acquire a specific bit of information in the domain. For example, a politician giving a speech from the stage is better positioned than the listeners sitting in the front row to see what is going on in the back of the room. Similarly, if I wanted to deposit a check in the bank and my brother was working just on the other side of the street, he might be better positioned to banish all my doubts about the closing time of the bank on the day of the town celebrations.

Cases in which someone serves the role of an epistemic authority for an interlocutor simply because she is better epistemically positioned than him constitute the minimal form of epistemic superiority. Let me stress once again that in cases like this, it is far from obvious that A is more reliable or epistemically virtuous than S: this explains why it is key to the plausibility of such a minimal account of epistemic authority that A be at least sensitive to the epistemic needs of her interlocutor. However, to the extent that S is conscientious enough to judge that, no matter how reliable or virtuous A is in general, A is better epistemically positioned than S is to get the truth in these specific circumstances, A's testimony still provides S with preemptive reasons to accept what she says.

Needless to say, cases in which A's epistemic superiority to S goes beyond mere epistemic position include situations in which A has more true beliefs and/or better understanding than S in a given domain. However, as I have tried to argue in this chapter, it is not enough for one to serve the function of an epistemic authority that they merely fare better than their interlocutor

[50] For similar considerations about the notion of *purely positional advantage*, see Williams (2002: 42) and Fricker (2006: 234-237).

as regards achieving epistemic goods such as truth, knowledge, or understanding. For a subject A^1 who is just a bit more competent in a domain than their interlocutor S but possesses and deploys the appropriate intellectual virtues can fulfill S's epistemic needs as good as—and, potentially, better than—a subject A^2 who has better competence in the same domain but displays a more limited range—or a less developed set—of novice-oriented abilities than A^1.

An interesting advantage of the framework that I am proposing is that it is compatible with the idea that there may be mixed cases of authority, such as the ornithologist who satisfies the requirements for being both an expert and an AofB and the professor who can be legitimately considered an expert and an AofU. The framework's adaptability and comprehensiveness are extremely crucial features for a theory that has the ambition to account for the various forms of epistemic authority, as it should now be clear that an epistemic subject might develop and refine their own research-oriented and/or novice-oriented abilities with time, that is, through exercise. In fact, we would wish that our epistemic authorities get better and better at fulfilling both the research-oriented and the novice-oriented function. These considerations speak in favor of conceiving the proposed theory of epistemic authority along the model offered by set theory. Besides allowing us to make sense of the intersections between the different sets, i.e., mixed forms of authority, it also gives us the possibility to move the various items or members from one set to another depending on the development of one's abilities and hence function.

The framework is organized around three main sets that capture the three forms of authority I have proposed (Experts, AofB, and AofU) and not most of the toy examples I have introduced so far in the thesis. Here are the three main sets:

[Expert] includes figures such as Shyler, whose superior reliability (or track record) in a given domain distinguishes them from the majority of epistemic subjects.

[AofB] includes figures such as my sister (from RED CARDINAL SISTER AGREEMENT CASE), who care that their interlocutor acquires true beliefs that they lack, by offering them preemptive or weighing reasons for adopting such beliefs.

[AofU] includes figures such as the grandmother (from GRANDMOTHER FISH CASE), advisors (as in Lackey's view) and, in general, those who care about helping their interlocutor improve their understanding of a

subject matter or a domain, by offering them preemptive or weighing reasons for adopting such beliefs.

As regards mixed forms of authority, the model can account for cases like the following:

[Expert + AofB] includes the ornithologist (from RED CARDINAL DISAGREEMENT CASE and RED CARDINAL AGREEMENT CASE), the doctor who tells me to cure my acid reflux rather than my cough and, in general, those experts who are willing to establish a relationship with the novices and care that their interlocutors acquire true beliefs in a given domain, by offering them preemptive or weighing reasons for adopting such beliefs.

[Expert + AofU] includes the professor (from the philosophy professor case as well as HISTORY PROFESSOR CASE) and, in general, those experts who are willing to establish a relationship with the novices and care about helping their interlocutor improve their understanding of a subject matter or a domain, by offering them preemptive or weighing reasons for adopting such beliefs.

[AofB + AofU] includes the grandmother (from GRANDMOTHER BOAT CASE) and, in general, those figures who care that their interlocutor acquires true beliefs and improves their understanding of a subject matter or a domain, by offering them preemptive or weighing reasons for adopting such beliefs.

Some might wonder why I have not yet placed Jäger's notion of Socratic authority in my framework. This is because to find the aptest position for it, we need to know whether he considers SAs to be not only AofU, but also experts according to the subject-independent or non-comparative dimension of expertise.[51] If the latter, then we might want to claim that SAs

[51] Notice that SA's having a higher ratio of true to false beliefs in a given domain than S merely allows us to consider her an authority of belief. In order for SA to be an expert as well, Jäger should define SA as having a higher ratio of true to false beliefs in a given domain than *most people do*.

represent the ideal sub-set of epistemic authority, one that fulfills all the requirements of the three notions above. If the former, we could simply consider SAs as both AofB and AofU, as they have both a "higher ratio of true to false beliefs in the domain than the subject does"[52] and novice-oriented abilities, maieutic ability included.

I take it that the ideal exemplar of epistemic authority has to display a further virtue, namely that of *wisdom*. Wisdom is the supreme ability possessed by those who recognize when it is better to directly settle the matter for S or to guide him in a more sophisticated manner by using maieutic techniques and further virtues.[53] It seems that an epistemic authority who is able to adopt both stances towards the layperson is epistemically superior even to Socratic authorities that always prevent themselves from settling the matter for S, since there might be cases in which doing so is more epistemically beneficial to S than just fostering his understanding. The GRANDMOTHER BOAT CASE and the GRANDMOTHER FISH CASE, taken together, feature an exemplar of epistemic authority (i.e., the grandmother) that is able to switch her attitude toward the layperson (i.e., her grandson) in such a way that she exemplifies the virtue of wisdom. The ideal exemplar of epistemic authority combines these features with objective expertise in a given domain, thereby fulfilling the three roles at once—[Expert] + [AofB] + [AofU]. This ideal epistemic authority might be defined as follows:

Ideal Epistemic Authority (IEA). A subject A is a IEA (for a subject S) iff

(1) A has more accurate information than the majority of people do in D;

(2) A possesses and makes extensive use of *research-oriented abilities*;

(3) A possesses and makes extensive use of *novice-oriented abilities*;

(4) A possesses and makes use of *wisdom*.

As a final remark, I shall briefly go back to the problem of expert testimony based on isolated second-hand knowledge that I have discussed in the previous section and highlight a further respect in which my theory of epistemic authority fares better than both purely preemptionist views and Lackey's advisor-as-authority model. In particular, given its pluralistic nature, my theory has two important advantages.

First, it allows us to grant the oncologist (see the example introduced in §3.5.3) the expertise she possesses for reasons that have nothing to do with the way she communicates to her patient that he has pancreatic cancer. (In fact, whether she has expertise entirely depends on

[52] See section 2.3 for Jäger's definition of SA.

[53] For further clarifications on wisdom and its various forms see, for instance, Ryan (2011), Whitcomb (2010), Zagzebski (1996).

the objective conditions outlined in chapter 2 and summarized early on in this section.) This is something that Lackey herself cannot account for because, as I have anticipated at the beginning of §3.5.3, she does not distinguish between experts and epistemic authorities.

Second, my theory allows us to calibrate the exact position of individuals such as the oncologist within the proposed framework of epistemic authority by analyzing the kind of service they provide. Lackey holds that someone who is not willing to address the questions raised by novices cannot be an epistemic authority, in that epistemic authority requires being able to advise the interlocutor. The view that I propose goes partly against this conclusion. It does remain true that a general unwillingness to address novices' questions undermines one's epistemic authority in the long run, because it constitutes a clear symptom of lack of sensitivity to the novice's epistemic needs. However, Lackey's theory, unlike mine, denies that some people can do us a great favor by passing along new information, despite the fact that they might lack the abilities that are required for addressing lay questions successfully. Clearly, we expect that oncologists—but, in general, physicians of any sort—can be both authorities of belief and authorities of understanding for their patients. Yet, there are other domains in which this is not strictly necessary: thus, it is an important value for a theory of epistemic authority that it can handle both kinds of situations.

In sum, the theory I have proposed here accommodates Lackey's idea that the capacity to advise an interlocutor constitutes a key dimension of being an epistemic authority—that is, something preemptionism typically fails to do—but it does not commit us to renounce another important function of epistemic authorities, namely that of providing novices with true beliefs they lack.

3.7 Conclusion

This chapter pursued a twofold aim. In the critical side of the chapter, I took issue with the Normative Question that lies at the grounds of this dissertation by offering a strategy to broaden Zagzebski's theory of epistemic authority in such a way that it can account for the authority of understanding, as well as by raising several critical points against various theories of epistemic authority on the market. First, I argued that Zagzebski and Jäger's accounts are affected by serious problems concerning the application of the Preemption Thesis and I provided reasons in favor of a limited version of the preemptionist view of epistemic authority. Then, I put forth compelling reasons to reject Jäger's view, in that it reduces epistemic authority to mere Socratic

exemplars, hence overlapping and underestimating the notions of an expert and of an authority of belief. Finally, I showed why Lackey's objections against preemptionist views of epistemic authority fail to hit their target and shed light on the limits of her advisor-as-authority alternative view.

In the constructive side of the chapter, I addressed the Conceptual Question that lies at the grounds of this thesis by endorsing an alternative framework of epistemic authority which sheds light on the intellectual virtues that each kind of authority is required to display. This virtue-based theory of authority is beneficial in three main ways: first, it overcomes several weaknesses in Jäger, Lackey, and Zagzebski's views; second, it yields a comprehensive and fine-grained framework that is in a position to account for many subtle differences among exemplars of epistemic authority; and finally, it vindicates the importance of distinguishing between the notion of an expert and that of authority in the epistemic realm.

Chapter 4

Epistemic Authority: A Collective Approach

Abstract. This chapter attempts to provide a remedy to a surprising lacuna in the current discussion in the epistemology of expertise, namely the lack of a theory accounting for the epistemic authority of collective agents. The strategy I pursue amounts to showing that the virtue-based theory of epistemic authority that I have proposed in previous chapters is well suited to account for the epistemic authority of collective bodies on a non-summativist perspective. More precisely, I show in detail how the defining requirements of an expert and an epistemic authority can apply to epistemic groups.

4.1 Introduction

As the previous chapters have shown, epistemologists working on the topics of cognitive expertise and epistemic authority are mostly concerned with establishing what it takes for an *individual* epistemic subject to be an expert or an epistemic authority, while few words have been spent on determining whether a *collective* agent such as an epistemic group or an institution can be an expert and what it takes for them to be such. Despite social epistemologists' broad interest in the epistemic agency of collectives, this issue is still in need of substantial consideration.

The aim of this chapter is to draw attention to this lacuna and offer a plausible account of what it takes for a collective agent to possess epistemic authority. Thus, the chapter purports to complete the answer to the Conceptual Question that I have addressed in previous chapters. Section 4.2 explores how the possession conditions of the proposed account of an expert—i.e., the epistemic superiority requirement and the (research-oriented) ability requirement—can apply to cases in which the expert is a collective body. Section 4.3 introduces the definition of a collective expert and illustrates some benefits of the proposed account. Finally, section 4.4 investigates how the account can be extended to the case of epistemic authorities.

Before we delve into the analysis, the project needs to be motivated. Why is it so important that we lay out an account that can accommodate the expertise of collective agents? The general motivation is grounded in the limits of mainstream individualistic epistemology, as in the Cartesian tradition that has dominated the epistemological discussion for centuries. As I have

mentioned in the introductory chapter of the dissertation, according to this tradition epistemology amounts to an inquiry into the "mental operations of cognitive agents in isolation or abstraction from other persons" (Goldman 1999, 4), where epistemic self-reliance is taken to be the main aim of epistemic agency. Fortunately, in the last decades this approach has been called into question by social epistemologists. As they rightly noticed, traditional epistemology fails to take into adequate consideration that our epistemic agency heavily relies on social sources of evidence and that collective entities may possess doxastic attitudes of the kind individual subjects possess (Goldman and Blanchard 2015).

This critique of Cartesian epistemology provides a specific motivation for a collective account of epistemic authority by pointing out that a substantial portion of our epistemic inquiry is conducted by collective agents of various sorts. This is obviously true of scientific and academic inquiry where most results are achieved by research teams and labs,[54] but it also applies to other epistemic endeavors such as information-gathering processes in journalism and assessments of legal evidence in court, where single professionals join their skills and epistemic resources together to achieve a joint epistemic goal. Since these days several information- and knowledge-producers are in fact collective agents, it becomes fundamental that we have an effective and reliable way to assess whether they possess epistemic authority. For if we do not know what it takes for them to be experts, we cannot evaluate which sources to trust.

4.2 Group Beliefs and Collective Virtues

The research-oriented view of an expert I defended in previous chapters accounts for the epistemic authority of *individuals* that fulfill the aforementioned requirements. My aim in this section is to show that the model can apply to *collective agents* such as epistemic groups, research teams, and professional units, thereby providing us with an account of what it takes for a group to be an expert. The model purports to do so in an elegant fashion, namely by showing that a collective agent can fulfill both requirements of the possession-condition of expertise, namely being epistemically superior to the majority of people in a domain—be it a matter of having more true beliefs or better understanding than most epistemic subjects do in

[54] Most accounts of the division of cognitive labor stress that competition among scientists is more effective than cooperation (see, e.g., Kitcher 1990; Strevens 2003; Zollman 2010). However, as Muldoon has recently argued (2017), a plausible account of the division of cognitive labor needs to explain the role scientific collaboration plays in current scientific endeavors, in that "the data suggest that collaborations are becoming increasingly integral to scientific production" (80). Arguably, the larger is the role of collaboration in science, the more fundamental it becomes to have an account of collective epistemic authority.

such domain—and possessing research-oriented abilities. I will assume that even collective agents have to fulfill RO-CAP, i.e., the functional definition of an expert on Goldman's view, as there is no reason why the function of an individual expert should differ from the function of a collective one.

To achieve the goal of this section, I shall accomplish two tasks: the first is to prove that some groups can be epistemically superior to others, while the second is to show that they can display intellectual virtues such as research-oriented abilities. The viability of the former task rests on important work on collective intentionality, group knowledge, and group understanding (e.g., Bird 2010; Boyd 2019; Gilbert 2013; Lackey 2014; Tuomela 1992), whereas to accomplish the second task I shall rely on Miranda Fricker's account of institutional virtues (2010).

4.2.1 The Epistemic Superiority of Collective Agents

Besides those who simply deny that collective agents can possess epistemic states—Boyd (2019) calls them *nihilists*—there are two main approaches one could take to account for the possession of epistemic states by collectives. According to a *deflationary* approach, a group G possesses the epistemic state X if and only if every individual member or most individual members of G possess X. This view is called deflationary or summativist (e.g., Lackey forthcoming) in that possession of the epistemic state X on the part of the group boils down to possession of X on the part of its members. To put it differently, the group comes to possess X as a result of summing up together members who possess X qua single individuals. The proposed definition remains neutral about the specific epistemic state that a collective agent acquires because it is meant to apply at least to the cases of belief, justification, knowledge and—as has been recently argued in Boyd (2019)—understanding. To give an example, on a deflationary account of group knowledge, we should contend that the sales department of Zara knows that the factory closes at 3:00 pm tomorrow if and only if most members of that department know about the early closure.

In contrast, those who feel the pull to hold that groups can possess or lack epistemic states despite what their members do qua single individuals will side with an *inflationary* approach. In this view, whether a group G possesses the epistemic state X does not depend solely on whether its members possess X qua single individuals (Boyd 2019: 3). The view is called inflationary or non-summativist in that possession of the epistemic state X on the part of the

group involves something that arises only at the collective level or, to put it differently, does not boil down to a mere sum of the epistemic states possessed by the individual members. Basically, any moderate version of the inflationary approach contemplates the idea that individual members still play a role in the process that results in the group possessing an epistemic state, be it belief, justification, knowledge, or understanding.[55]

A typical version of this view, due to Margaret Gilbert's work (e.g., 2013), requires that individual members, despite lacking individual knowledge that *p*, share a joint commitment to the proposition that *p*, or accept that *p*. For example, a group of engineers who are testing a new material to create drones can know that the material is sufficiently resilient even when every member does their own specific part of the test and inserts their results into a computer that autonomously connects all the information and makes a reliable prediction. On a moderate non-summativist reading of this case, the team as a whole knows that the new material meets the required standards of resilience, even though no single engineer possesses this knowledge.[56]

Having laid out the main approaches to collective epistemic states, we can start reflecting on their suitability for our purposes—that is, on whether we can apply them in setting up a collective epistemology of expertise. For the sake of simplicity, I shall begin by considering the case of group knowledge. Two reasons support this choice. First, although in the previous chapters I have argued in favor of an account of an expert that cashes out their epistemic superiority in terms of understanding, group understanding is a trickier matter than group knowledge, as I shall show later on in this section. Second, even though Goldman's original possession condition involves true beliefs rather than knowledge (see §2.2), he does so for reasons that go well beyond the scope of this chapter (see, e.g., 2001: 92) and the notion of group knowledge is close enough to serve my purposes here.

The ideal scenario is one in which we could remain neutral as to which account of group knowledge works better in our framework. However, when it comes to the epistemology of expertise, we cannot but point out that the inflationary approach has better prospects for integration with a theory of epistemic authority, in that this view is better placed to account for a fundamental principle of any scientific enterprise: the division of cognitive labor.

Deflationary views of group knowledge do not seem to be compatible with contemporary scientific teamwork because they are committed to maintaining that for a group such as a

[55] See Lackey (2014) for a version of *moderate non-summativism*. In contrast, *radical non-summativism* holds that a group G can know that *p* even when not a single member of G is aware that *p*. See Bird (2010) for a version of radical non-summativism, and Carter (2015) for a clear overview of the divide in the epistemology of groups.

[56] Another set of cases that meet the requirements of *moderate non-summativism* includes situations in which all or most individual members of a group G are aware that *p* yet do not believe it for some reason. For an example, see Carter (2015: 714, n.8).

research team of experts to know that *p*, most or all individual members of the team should know that *p*. But the organization of science—especially in interdisciplinary works—is such that more and more people are assigned specific tasks within a broad project while fewer and fewer people are in charge of analyzing the results and drawing the right conclusions. Hence, it is not surprising that at the time of the publication of a scientific team's results (*p*), only a few members of the team know that *p*. In contrast, an inflationary view like moderate non-summativism allows us to hold firm that the team as a body knows that *p* insofar as individual team members jointly commit to pursuing the project with its proposed aims and structures and to accepting the results of the study as conveyed by the people in charge.

Once plausible accounts of group knowledge have been introduced and discussed, it takes only a few steps to get to a collective version of the truth-linked condition for expertise. From an inflationist perspective, proving that a group can satisfy T-LCa becomes merely a matter of a comparative evaluation of the epistemic standings of various agents. In fact, all we need to do is show that collective entities can have more true beliefs than other agents in some domain. Specifically, a group G (e.g., a team of oncologists at the European Institute of Oncology) fulfills T-La in domain D (e.g., the domain of oncology) insofar as G has more true beliefs than most subjects and groups have in D.

However, as I argued in §2, a more promising account of an expert involves an understanding-linked version of the possession condition, according to which experts are required to possess better understanding than most people do in a domain. Social epistemologists have just started to inquire into whether collective agents can possess understanding.[57] Interestingly enough, Wilkenfield et al. (2016) have shown that laypeople use attributions of understanding to identify people to whom they are willing to defer—that is, people they take to be experts—as well as to distinguish these individuals from other potential informers. This consideration brings support to the overall project of grounding the notion of expertise on understanding rather than mere knowledge. It also supports the idea of extending the attribution of understanding to groups because it is not at all uncommon for laypeople to put their trust in a collective agent rather than a single expert. To give just an example, a couple might trust the medical team of a pediatric clinic to ensure that their unborn child receives all the therapies she needs to minimize the risks of premature death. Since the health conditions of the baby are extremely complex, the situation requires a multi-disciplinary understanding that only a team of medical experts can provide. For none of the team members—not even the

[57] To my knowledge, the only published works addressing this issue are Boyd (2019), Brady (2016), and Delarivière (2020).

pediatrician that coordinates the team—qua single expert possesses enough understanding and competence to treat the unborn baby as she needs.

It seems thus epistemically appropriate, but also practically plausible, to imagine that the couple trusts the team rather than their members or the coordinator. This is not to say that they will cease or fail to acknowledge that the team is constituted by a number of medical experts. But were something to go wrong, the entire team would be responsible for it: even if the couple complained to the coordinator, he or she would be the main subject of the complaint just because of his or her coordinating and spokesperson function within the team.[58]

In a similar vein, we could imagine that the coordinator of the medical team informs the couple about the recent developments of the baby's conditions and the therapies they recommend. It is not hard to suppose that the explanation offered by the doctor puts together (and simplifies as much as possible) various inputs offered by the team members based on their respective specialisms and the consequent joint decisions that they took as a team. Surely, the coordinator's novice-oriented abilities play a role in ensuring that the couple understands the medical situation of their unborn baby, but it seems fair to contend that the team itself is their source of understanding. Thus, the plausibility of the idea that groups can work as sources of understanding provides further reason for thinking that groups can possess understanding (Boyd 2019: 10).

However, when it comes to explaining what it takes for a collective body to have understanding, some conceptual machinery is required. The notion of understanding that I have outlined in chapter 3 includes an informational component as well as a grasping component. As should be clear, it is not hard to make sense of what it takes for a group to satisfy the former requirement. If we stick to the pediatric team example, for the team to fulfill the informational component it is enough that the information which the team members jointly commit to accepting encompasses what is necessary to possess understanding in the domain of pediatrics. The complexities arise when it comes to explaining how a group satisfies the grasping component, in that grasping involves a broad range of abilities. In particular, according to Hills, understanding entails that, in the appropriate circumstances, epistemic agents can successfully (2009: 102-103):

(i) Follow an explanation of why *p* given by someone else;

[58] In the remainder of the section, I will work with the example of the group coordinator rather than a proper authorized spokesperson, who might or might not be a member of a group depending on the specifics of the group itself. More on the role of authorized spokespersons in group epistemology can be found in Lackey (forthcoming: §4).

(ii) Explain why *p* in their own words;

(iii) Draw the conclusion that *p* (or that probably *p*) from the information that *q*;

(iv) Draw the conclusion that *p'* (or that probably *p'*) from the information that *q'* (where *p'* and *q'* are similar to but not identical to *p* and *q*);

(v) Given the information that *p*, give the right explanation, *q*;

(vi) Given the information that *p'*, give the right explanation, *q'*.

These abilities describe what it takes for someone to have what Hills calls *cognitive control* over a proposition (2016: 663). At the objectual level, it is reasonable to assume that displaying understanding of a domain requires extending such a local cognitive control to encompass a comprehensive grasp of how the various pieces of information relate to each other and to other elements of the domain to constitute a coherent account of the matter (Elgin 2017: 46). In sum, grasping requires that a group be able to manipulate available information in certain ways and produce some kind of explanation of how the various pieces of information are related to each other in a given domain.

Before going on with the argument in support of a notion of group understanding, it is important to get clear on one aspect of grasping, namely the relationship between having understanding and possessing the ability to provide relevant explanations of propositions within a given domain. At first glance, it might seem as though Hills' account of grasping clashes with the account of an expert that I have defended in previous chapters. For, as I argued with the case of Shyler (§2.5), someone can serve the function of an expert without being willing and able to address the novices' epistemic needs and this seems inconsistent with requiring that experts possess the ability to provide relevant explanations within their domain of expertise.

On closer inspection, though, my theory of epistemic authority can appropriately discriminate between two different ways in which someone can provide the right explanation to some information or query. What is required of experts is that they know and understand the relationships between various pieces of information that their area of expertise encompasses and, in turn, that they can make reference to such relationships while explaining why it is the case that *p*. This is a necessary component of grasping, namely one that experts are supposed to fulfill (even) on my account.

In contrast, my account does not ask of experts that they have the ability to fulfill a more complex function, namely that of providing laypeople with the explanations that are appropriate to their level of competence within the domain. In other words, I take it that the grasping component of understanding—both at an individual and a collective level—does not include

what I have previously defined as novice-oriented abilities. Evidently, we would be willing to say that someone has understanding of some domain if they are able to clarify issues within such a domain for the benefit of laypeople. However, it is compatible with possessing understanding in a domain that one, for example, fails to acknowledge why it is so difficult for laypeople to figure out what is going on within that domain and why information *p* and *q* are obviously related in the way that they are. Similarly, we would not want to deny that someone has understanding in a domain merely because they lack the ability to express themselves in a way that is perfectly accessible to novices. These considerations explain why from the fact that experts need to possess cognitive control within their domain of expertise, it does not follow that they also have to possess novice-oriented abilities.

That said, it is still an open question how a collective body can fulfill the grasping component of understanding. Boyd's recent work offers the following account of epistemic grasping at a collective level:

> *Group Grasping*: A group G grasps *p* and its relationship to reasons that support *p* just in case (i) G represents *p* and reasons for *p*, and (ii) the members of G are *mutually p-reliant* (2019: 16).

The first condition of this definition relies on the idea that any agent—be it an epistemic subject or a collective body—has to entertain some sort of mental representation of the object that they understand. This condition appeals to Wilkenfeld's representation manipulability account of understanding (2013) and has been extensively discussed in Kelp (2015). Boyd (2019: 6, 16 n.22) suggests that a promising form of mental representation in the case of understanding is belief, but remains neutral on what the best theory of group representation is. For the sake of this chapter, I shall set aside further discussion on this issue and focus on the second condition of the definition of group grasping.

According to Boyd, for a group to grasp a proposition *p* it is necessary that the members of the group "recognize both that they are contributing towards the relevant goal (perhaps in the form of representing reasons and relationships between reasons), and that they would not be able to achieve that goal on their own" (17). To illustrate this with an example, consider again the pediatric team and imagine that its members, as a matter of fact, do not work as we would expect of a team. Each specialist does their own part, but they tend to double-check the contributions of other members as much as possible or consult external experts without letting the others know because they lack trust in each other. If everything goes well, they end up

issuing the same diagnoses and propose the same therapies than more virtuous groups, but they do so in a confused and random fashion. By contrast, a mutually-p reliant medical team is composed of members who entrust that everyone does their own part to the best of their knowledge and share the joint goal of taking care of their patients in the most effective manner, with the awareness that none of them as a single medical expert would be able to get a grip on the patients they are treating and envision an effective therapy for them. For these reasons, we should expect of this group that it grasps the complexity of the medical cases it is requested to handle and the prospects of the therapies it proposes to the patients. This explains why in the latter case, unlike the former, we could contend that the team, qua collective agent, possesses group understanding according to an inflationary perspective.

It is worth stressing again that the view outlined here is the first attempt in the literature to account for what it takes for collective agents to possess understanding. It can be expected that the debate around this issue will grow in the future and more accounts of group understanding will be proposed. Such a scenario should be welcomed by epistemologists of expertise as well, in that it would provide the resources to go beyond an account of the epistemic superiority of collective agents whose main contours I have just delineated here with the available conceptual tools.

4.2.2 Collective Agents and Their Intellectual Virtues

Let us move now to the second part of the job: to illustrate how collective agents can possess intellectual virtues and, in particular, research-oriented abilities. To accomplish this task, I shall rely on Fricker's (2010) twofold model of group virtues, which applies Gilbert's notion of joint commitment to group *motives* and group *ends*.[59] Fricker introduces the distinction between motives and ends because she wants to highlight that her model is suitable for both *motive-based* virtues and *skill-based* virtues.

As an example of the former, she mentions kindness, compassion, charity, and generosity. For an agent to possess motive-based virtues, they need to fulfill two requirements. First, they need to commit to achieving the good end of the motive for the right reason, namely by acknowledging that it is good to have the motive which, other things being equal, will bring about further good. Second, they need to reliably exercise the virtue in the appropriate circumstances. For example, in order for one to display kindness, it is not sufficient that they

[59] For an alternative take on group virtue epistemology, see Kallestrup (2016).

help the first old passer-by cross the road. Rather, one is authentically kind insofar as they reliably adopt this attitude out of a commitment to the motive of kindness as they judge that it is intrinsically good to have this trait. For the value of motive-based virtues lies in possessing their good motives, not in the further goods they bring about.

In contrast, the latter type of virtues includes traits such as vigilance, honesty, justice, and inventiveness that an agent can naturally possess without possessing virtuous motives, insofar as, first, they commit to bringing about the virtuous end; and second, they reliably exercise the virtue in the appropriate circumstances, no matter how conscious of its intrinsic value they are. For example, for a soldier to possess the virtue of vigilance, she needs to commit to achieving the end of vigilance through a reliable method—by, say, being on the alert for enemy movement—even though this attitude is somehow part of her background military training and she has never reflected on its good outcomes. For the value of skill-based virtues merely lies in bringing about their good ends.

Let us now analyze how Fricker's model accounts for motive-based virtues and skill-based virtues when the virtuous agent is a collective. On the one hand, what allows a group to display a motive-based virtue is the fact that (a) its members jointly commit to the *motive* of that virtue and (b) the group proves to reliably achieve this end. As regards requirement (a), a joint commitment is in place if all individual members express their "readiness to be jointly committed to espouse the relevant goal [the motive of the virtue] as a body" (Gilbert 2013, 32).

Consider a relevant case for the purposes of this chapter. Imagine an expert team of biologists who jointly commit to a research-oriented ability such as open-mindedness. Fricker's view grants that the team is open-minded insofar as its members have the appropriate motivation, that is, if they are willing to join forces in adopting an open-minded behavior because they acknowledge that the trait is good per se and hence displaying it would contribute to the flourishing of the group and the scientific community. Requirement (b) captures a widely shared idea among virtue theorists according to which an agent's character trait amounts to a virtue insofar as the agent is reliably successful in bringing about the specific end of the virtue.[60] The team of biologists fulfills this condition insofar as it proves to be open-minded in its research and agency by, for example, refraining from dismissing opposing theories and studies, being willing to reconsider their theories in light of counterevidence, and counteracting confirmation bias and similar forms of prejudice.

[60] See, in particular, Zagzebski's account of a virtue as involving a motivational component and a success component (1996, §2).

It is important to notice that the team might display open-mindedness even if its members, taken individually or outside their professional activity, lack the virtuous trait. On Fricker's non-summativist or collectivist account, this can be explained by pointing out that each of us has various practical identities, which can sometimes be in tension with each other. Thus, it might be the case that a subject has a virtue only as a member of some group and not as a private individual (2010, 238).

On the other hand, Fricker's account of a group's skill-based virtues partly differs from the previous one in that they share condition (b) but not condition (a). In fact, a group possesses a skill-based virtue insofar as (a*) its members jointly commit to the *end* of that virtue and (b) the group proves to reliably achieve this end. Let us apply the account to another case of a research-oriented ability. Consider again the team of biologists and suppose all members jointly commit to taking a public stand against anti-vax movements and defending the value of their research despite its unwelcome implications for the current government's political agenda. In such a case, the absence of a joint commitment to the virtuous motive of intellectual courage is replaced by the team's pooling their faculties and jointly committing to bringing about the virtuous end of intellectual courage through the division of labor, as required by Fricker (243). For example, we can suppose that some members take care of drafting a public statement meant to explain the value of the team's research to a wide audience, others discuss how to respond to political attacks, and a third group engages with the institution's administrative board to look for their support in the public debate.

Based on these considerations, I think we can agree with Fricker that the team of biologists displays intellectual courage even though its members lack full awareness that they individually possess this trait and the appropriate motivation.[61] For, as we have already seen in the soldiers case, an agent—here, a collective agent—can develop skill-based virtues despite a lack of motive insofar as their performance reliably brings about the virtuous end.

The two cases of the team of biologists are examples of how a collective agent can display relevant intellectual virtues for the service conception of epistemic authority. However, this view of collective intellectual virtues brings with it at least two problems we need to address here. The first issue concerns how research-oriented abilities map onto Fricker's twofold account of virtue. I am inclined to think that this set of intellectual virtues can either develop as motive-based virtues or as skill-based ones depending on the specifics of the case in question. Several virtue theorists would resist the idea that one can be, for example, intellectually courageous while lacking the respective motivational state, as we commonly conceive

[61] See Lahroodi (2007) for further considerations on members' awareness in relation to group virtues.

intellectual courage as a virtue of intellectual character that arises out of an agent's love of truth.[62] Nonetheless, I shall grant that an expert group can display this virtue simply because its members jointly commit to performing their task but divide the labor in such a way that their agency reflects the group's trait.[63]

Let us set aside this issue and consider an objection against the collectivist spirit of Fricker's account recently raised by Sean Cordell (2017). Specifically, the objection tackles the idea that joint commitments to a motive or an end of a virtue are irreducible to personal commitments. Cordell accepts that joint commitments involve "hat-wearing" members who have different practical identities and therefore act as they do in light of their roles as members of a group (e.g., the biologists jointly commit to open-mindedness qua members of the team, yet they might individually lack this trait when acting alone or in other spheres of their agency). However, he argues that joint commitments boil down to individuals' commitments because in fact "it is individuals who make their group-oriented joint commitments albeit crucially, and only, when wearing their 'hats' which comprise their institutional social role and its obligations" (Cordell 2017, 48).

I shall resist Cordell's objection by arguing that the mere fact that group virtues involve individuals' commitments does not undermine the irreducibility of group virtues to individual ones; that is, it does not reduce Fricker's account to a summativist view of collective virtues. For it should be taken for granted that all members of the group have to do their own part—that is, individually commit to the motive or the end of a virtue—in order for the group to display the collective virtue. The same is true of non-summativist accounts of group knowledge where, as the example of the team of engineers shows (see §4.2.1), each member's contribution to the test makes it possible that the group as a whole acquires knowledge about the new material's resilience. But much as, in the engineers case, group knowledge does not boil down to each member's work on a part of the test because each member may in fact lack knowledge of the result of the test, so, in the biologists case, group virtue does not boil down to each biologist's individual commitment to the motive of open-mindedness (or the end of intellectual courage) because each biologist may in fact lack open-mindedness (or intellectual courage).

Furthermore, as Gilbert pointed out, individual commitments on the part of group members are best intended as *dependent individual commitments* that (i) only exist through the joint

[62] See, e.g., Annas (2011) and Zagzebski (1996). In contrast, Driver (2013) and Stichter (2018) endorse a model of virtues as practical skills that captures some features of the proposed account. See also Chi (2006, 24) and Dreyfus-Dreyfus (1991) for psychological evidence supporting the thesis that possession of a virtue does not always require the capacity to articulate one's reasons.

[63] See, e.g., Pritchard (2018) for a clear examination of the distinction between motive-based intellectual virtues and skill-based cognitive faculties.

commitment, (ii) aim at promoting its object, and (iii) are *simultaneous*—that is, they "come into being simultaneously at the time of the creation of the joint commitment" (2013, 41). Thus, although it is true that for a group to possess a collective virtue something is required from each individual member, that is not enough to reject a non-summativist account of collective virtue.[64]

4.3 Collective Experts

In the last section, I showed how possession of research-oriented abilities can be attributed to a collective agent based on Fricker's non-summativist account of group virtues. Before putting together the proposed conditions of collective epistemic authority, I shall point out that the plausibility of my account should not become hostage to a non-summativist model of group virtues. In fact, it is compatible with the functional definition of an expert introduced in chapter 2 (i.e., RO-CAP) that a collective agent possesses a virtue insofar as each member individually possesses that virtue. Thus, those who are not convinced by my arguments in favor of a non-summativist approach can still find scope for an account of collective epistemic authority in the service conception.

Let us combine the conditions proposed in the last sections and outline the service conception on a collective level:

> *Collective Expert-B*. A group G is a collective expert-b in domain D if and only if (i) G has more true beliefs than most agents have in D and (ii) G possesses research-oriented abilities.

> *Collective Expert-U*. A group G is a collective expert-u in domain D if and only if (i*) G has better understanding than most agents have in D and (ii) G possesses research-oriented abilities.

As should be evident, condition (i) amounts to a collective version of T-LCa, while condition (i*) amounts to a collective version of U-LC. I endorse a non-summativist reading of both versions of the possession condition of expertise because, as I argued in §4.2.1, the dynamics of scientific research are such that a group as a whole can retain knowledge that *p*

[64] For another worry against Fricker's collectivist account of group virtue, see Cordell (2017, 48-50).

when its members in fact lack knowledge that *p*. Condition (ii) extends to collective bodies the requirement that an expert needs to possess research-oriented intellectual virtues.

On the proposed non-summativist reading, a group can possess these virtues even if its members, taken individually, fail to have them. This allows us to explain why, say, a scientific team is intellectually curious or open-minded even though some of its members lack such traits. In other words, it allows us to assess whether a collective agent possesses a virtue based on the group's attitudes and actions as a whole, no matter whether its members, taken individually, possess the virtue. Those who endorse a summativist view of group virtues can read (ii) along their favored lines and contend that a group possesses research-oriented abilities insofar as all (or most) members of the group possess those intellectual virtues and deploy them in joint activities.

Adopting a service conception of epistemic authority to define what it takes for one to be a collective expert can improve our understanding of cognitive expertise in several ways. First, it allows us to shed light on the weakness of a collective version of a purely veritistic, or truth-linked, account of expertise such as Coady's (2012), according to which fulfilling a truth-linked requirement such as condition (i) suffices to attribute collective expertise to a group. The scope of such an account would be very limited, for it would at best allow us to flag and possibly isolate fake experts or unreliable sources of information. Let's be clear: this is a fundamental feature of an account of cognitive expertise, especially given the widespread amount of misinformation and fake news we encounter on the media and social networks. However, we should aim for something more.

By requiring that collective experts need to possess research-oriented abilities, the service conception of epistemic authority provides us with the resources to identify those groups—for example, teams of scientists or labs—that have more prospects to make a contribution to the epistemic progress of a discipline through their research. This feature of the service conception is not only epistemologically but also socially and politically significant, in that it puts policy makers in a position to decide which kind of groups need to be supported more depending on whether they want to foster institutions that have the capacity to innovate or institutions that can help laypeople improve their epistemic competence in some domain.[65]

Finally, including group virtues in the definition of a collective expert could provide groups that aim at fostering the epistemic progress of our communities with directions for ameliorating their agency. A genuine inquiry into the abilities that experts need to possess

[65] In other words, the social significance of distinguishing between groups that fulfill a research-oriented function and groups that fulfill a novice-oriented function is to offer a theoretical justification for implementing strategies to measure the quality of research and teaching of various institutions at a national and international level.

depending on the function they are supposed to fulfill would in fact allow them to contrast their activity with the ideal requirements of a virtuous collective expert and evaluate how they could better perform their social and epistemic role. For as Goldman points out, "if we wish to raise our intellectual performance, it behooves us to identify those traits which are most in need of improvement" (1978: 511), and there is no reason why this claim should not be also true of collective agents.[66]

4.4 Collective Epistemic Authorities

Before concluding, I shall briefly inquire into another important dimension of what I would like to call *the epistemology of collective expertise*, namely the epistemic superiority of collective epistemic authorities. The argument of this chapter focused on the plausibility of extending the attribution of expertise to collective agents. The reader might rightly wonder what I would have to say about epistemic authorities, given that on my theory they constitute a different form of epistemically superior agents than experts.

One reason for thinking that groups can serve the function of epistemic authorities is that the conceptual structure of the notion of an epistemic authority is very similar to the definition of an expert and I have already shown that groups can possess expertise. As argued in §3.6, for someone to be an epistemic authority it is required that they are better epistemically positioned than their interlocutors in a given domain, where being in a better epistemic position results in a (perhaps limited) superiority in terms of true beliefs and/or understanding. The considerations offered in §4.2.1 illustrate that extending the superiority requirement of epistemic authorities to collective agents would not be a problem. Epistemic authorities are also required to possess novice-oriented abilities, that is, those intellectual virtues that allow them to address the interlocutor's epistemic needs in a given domain. Without going into the details, it should be evident that this requirement too could be extended to account for collective epistemic authorities, in that the argument offered in §4.2.2 specifies how possession of intellectual virtues can be ascribed to collective agents.

What could constitute an obstacle to an account of a collective epistemic authority is the remaining condition in the proposed definition, namely the fact that epistemic authorities are

[66] Recent work in vice epistemology (e.g., Cassam 2016) could contribute to the understanding of expertise in the opposite way—i.e., by shedding light on the intellectual vices that prevent a group or an institution from becoming an expert collective. See, e.g., Lefevere and Schliesser (2014) for considerations on collective negligence and responsibility in science.

meant to have an ongoing relationship with their interlocutor. Now, as I pointed out in §3.2, the reason why epistemic authorities need to have a relationship with the subject who defers to them is that a relationship activates their sensitivity to the novice's epistemic needs, thereby enabling them to address the queries of the interlocutor in a way that suits her epistemic possibilities. Some might suspect that this condition cannot be met at a collective level, in that it is hard to imagine how a group can display a sensitivity that is triggered by particular features of a conversation between epistemic subjects like the implicatures that both parties deploy, the emotional clues they convey, and so on.

Going back to the example of the multi-disciplinary pediatric team, it might seem as though sensitivity to the epistemic needs of the parents of a young patient can best be displayed by the coordinator of the team, who is in charge of keeping them updated on the medical conditions of their kid and the therapies that the team considers to be most appropriate to treat him. As a matter of fact, the coordinator is the member of the team that has direct contact with the parents and therefore can display the required sensitivity to their epistemic needs.

In the remainder of this section, I shall attempt to sketch a reply to this objection here, which sheds light on a key problem for my theory of epistemic authority. It is important to point out from the start that the following considerations are only meant to set out what I take to be a promising line of response—but still one that is in need of further argumentative support. I think it is worth flagging the problem although the currently available resources in group epistemology—in particular, in group virtue epistemology—might not allow us to settle the matter completely.

Let me start with what the objection gets right, that is, the plausible idea according to which, in some cases, a selected portion of group members—potentially, only the coordinator—has to exercise a great deal of the sensitivity required for the group to fulfill the service of an epistemic authority. But what the objection gets wrong is the idea that such an imbalanced exercise of an epistemic virtue among the group members will result in the group lacking the virtue —or, in the virtue not being possessed by the group itself as opposed to the relevant members.

One way to resist the objection is to argue that for a collective agent to serve the function of an epistemic authority it is merely required that the relevant novice-oriented virtues are possessed and exercised by its members, no matter how many of them deploy a given ability and to what degree each of them exerts it. For all that matters is that the group members jointly commit to the appropriate division of epistemic labor (as in Fricker's skill-based account of group virtues) and each member does their own part. On the most charitable reading, this view commits us to conceding that it is enough for, say, the pediatric team to fulfill the dispositional

condition of epistemic authority that the coordinator alone displays the required sensitivity to the epistemic needs of the patient's parents.

If this is the case, then this view seems to face a structural asymmetry: it relies on an inflationary account of the possession condition of a collective epistemic authority—no matter whether we cash out the epistemic superiority in terms of true beliefs, knowledge, or understanding—as well as on a (particular form of a) deflationary account of the dispositional condition, which is fulfilled to the extent that the group's coordinator exercises the required sensitivity to the epistemic needs of the interlocutor. And it looks as though the only available way to motivate this asymmetry is to invoke some sort of structural difference between individual and collective epistemic authority and, in turn, divide the epistemic labor in such a way that the group's coordinator is responsible for understanding the epistemic needs of the interlocutors and tailoring the group's epistemic services to their epistemic resources.

Despite the plausibility of the idea that the coordinator has some specific responsibilities that other group members do not share, this solution to the above objection might still appear a bit *ad hoc*. In the end, if it is just the coordinator who has to display novice-oriented abilities, then it is hard to see how these virtues can be attributed to the group as a whole. However, I think that the problem of the collective version of the sensitivity to the epistemic needs of the novice can be handled in a different way, namely one that is more faithful to the inflationary spirit of the theory of collective epistemic authority that I have proposed in this chapter.

The crux here is to notice that the role of the coordinator is less decisive than I have attempted to argue so far in this section. Consider again the example of the pediatric team. It is certainly true that the coordinator has to be sensitive to the epistemic needs and the epistemic resources of the patient's parents, but it is not the case that the epistemic labor of the team can be divided in such a way that she is the only member who has to exercise these intellectual virtues. If this were the case, it would become more and more complicated for the group to fulfill their service. We can imagine that, in practice, this sensitivity involves things like selecting and putting together those (and only those!) relevant diagnostic data about the kid's conditions that the parents can understand; explaining the pros and cons of different surgical approaches that might be undertaken; illustrating the long-term medical therapies that the kid might have to undergo depending on the development of his situation; and so on.

It might be the case that the group coordinator could do all this work on her own, but a more plausible view of the division of epistemic labor in the medical team also involves the members' joint commitment to bringing about the motives or the ends of the sensitivity in question. Practically speaking, the commitment might involve a mutual acknowledgment of the

motives of the novice-oriented ability in question—which would allow us to consider the sensitivity as a motive-based collective virtue. Alternatively, it might merely lead all the units of the team to share with the coordinator some notes about how to present the situation to the patient's family or reserve a specific portion of their meetings to discussing with her how best to manage the relationship with the parents in light of the emotional reactions they have exhibited so far—which would allow us to consider the sensitivity as a skill-based collective virtue.

In sum, although the group member in charge of dealing with the novice cannot lack the sensitivity to the novice's epistemic needs and resources, it seems as though the most effective way for the group to fulfill the function of an epistemic authority involves more than a division of the epistemic labor on which the group members divide among each other the novice-oriented abilities that they are supposed to exercise. As the case of the pediatric team shows, the group has better prospects to meet the epistemic needs of the interlocutor—or, at least, to do so in a more efficient and efficacious way—if all its members jointly commit to exerting the required novice-oriented abilities and cooperate towards bringing about the motives or the ends of such virtues.

As stated at the beginning of this section, these considerations on the notion of a collective epistemic authority were primarily meant to broaden the discussion in the epistemology of collective expertise by showing how it could be possible to account for the epistemic authority of group agents without losing the benefits of a virtue-based framework. More needs to be done to provide a full epistemology of collective expertise as well as to explain how collectives can possess (epistemic) virtues.

4.5 Conclusion

Let me take stock. In this chapter, I tried to outline a collective account of expertise based on what I have called a service conception of epistemic authority. I have done so by providing a collective version of the possession condition of expertise—in its truth-linked and understanding-linked forms (§4.2.1)—and by showing that we can demand that collective experts possess research-oriented abilities since groups can have virtues (§4.2.2). I also shed light on some epistemic and social benefits of the proposed account of collective expertise (§4.3). Finally, I suggested how this analysis could be extended to provide an account of collective epistemic authorities and discussed a relevant problem with the possibility to ascribe

the sensitivity to the novice's epistemic needs to group agents (§4.4). The main goal of this chapter was to motivate the need for inquiring into the nature of group expertise and to suggest a trajectory for filling a significant lacuna—that is, the absence of an account of a collective expert in the epistemological literature. By showing how my epistemology of expertise applies to the case of collective bodies, this chapter has completed the answer to the Conceptual Question that lies at the ground of this thesis. In the remainder of this dissertation, I shall devote my attention to the Normative and the Practical questions.

Chapter 5

Epistemic Paternalism and the Service Conception of Epistemic Authority

Abstract. Epistemic paternalism is the thesis that in some circumstances we are justified in interfering with the inquiry of another for their own epistemic good without consulting them on the issue. In this chapter, I address the issue of who is rationally entitled to undertake paternalistic interferences, and in virtue of which features one has this entitlement. First, I undermine the view according to which experts are the most apt people to act as paternalist interferers. Then, I argue that epistemic authorities are in a better position to satisfy the requirements of justified epistemic paternalism, when conceived according to the service model of epistemic authority. Finally, I offer a virtue-based account of paternalist interferers and show how it can apply to cases in which the interferer is a group or an institution.

5.1 Introduction

Suppose a mother enrolls her son in university and pays for the tuition fees while he is working to save money and is unsure whether to keep studying. Almost everyone would agree that she is acting paternalistically towards him. Some would argue that her interference is permissible; others would disagree. In this chapter, I am interested in the specific phenomenon epistemologists call *epistemic paternalism*, according to which in some circumstances we are justified in interfering with the inquiry of another for their own epistemic good without consulting them on the issue (Ahlstrom-Vij 2013: 4). For now, let us consider the example above as a case of an epistemically paternalistic interference. Other examples may include a judge withholding information from the jurors about the track record of crimes committed by the defendant in order to preclude them from developing a bias against him; and a health department mandating the introduction of prediction models for medical diagnosis and prognosis in order to prevent clinicians from overestimating their expertise and clinical abilities.

Epistemic paternalism is commonly regarded as a harmful epistemic practice that could undermine our freedom, epistemic autonomy, or both. However, in the last three decades, a few epistemologists have endorsed the view that there are both genuinely defensible forms of epistemic paternalism and epistemic goods that paternalistic interferences could allow the

subjects interfered with to gain (Ahlstrom-Vij 2013; Bullock 2018; Goldman 1991; and Pritchard 2013b). Surprisingly enough, not much work has been done on the question of who is rationally entitled to undertake paternalistic practices, and in virtue of which features one has this entitlement.

I aim to provide a compelling answer to this question as a way to address one key aspect of the rational boundaries of the interaction between non-peer agents. Thus, this chapter contributes to outlining an answer to the more general Normative Question that lies at the grounds of this thesis. In particular, I will challenge Goldman's view, according to which one's paternalistic interference is justified insofar as the interfering subject is an *expert*. After reconstructing the core thesis of epistemic paternalism and discussing the main questions it generates (§5.2), I shall argue that the epistemic conditions for being a paternalist interferer substantially differ from the requirements of cognitive expertise (§5.3). Specifically, they differ in a way that makes Goldman's own definition of an expert inadequate to justify epistemically paternalistic interferences, as paternalist interferers have a different task to accomplish from experts and therefore are required to display a different set of intellectual virtues. I shall also argue that *epistemic authorities* cannot fulfill the function of paternalist interferers (§5.4). Yet I will show that experts and epistemic authorities have some relevant features in common. In §5.5, I will offer what I consider a compelling account of virtuous paternalist interferers, while in §5.6 I shall defend the idea that my account can apply to cases in which paternalist interferers are collectives, such as groups or institutions.

Methodologically speaking, the argumentative strategy I shall pursue in this chapter applies the virtue-based approach to the epistemology of expertise that I have developed in the previous chapter to the issue of epistemic paternalism. This analysis would bring further support to the idea that my theory of epistemic authority provides an extremely effective tool for distinguishing various ways in which a subject can be epistemically superior to another. Furthermore, the argument of this chapter shows how virtue theory contributes to the current epistemological research by providing insights into an underexplored topic in social and applied epistemology.

Some might feel disappointed about the scope of the project in that it does not purport to provide a conclusive straightforward answer to whether epistemic paternalism, in general, is an epistemically justified practice. That remains a fair question, one that I will attempt to address in the following chapter, but I am convinced that research on epistemic paternalism could nonetheless benefit from the results of the project I shall pursue in this chapter. Were my argument to be compelling, it would provide an effective corrective to a potentially wrong

research line according to which only experts should be granted the entitlement to paternalistically interfere with someone's inquiry. It would also allow us to identify another type of authoritative subject in the epistemic realm, one that should not be confused with cognitive experts and epistemic authorities.

5.2 Epistemic Paternalism in a Nutshell

In Alvin Goldman's early formulation (1991: 118-119), epistemic paternalism has two fundamental features. First, it is a form of *protection* that a subject (or a group) A, who is more reliable than a subject (or a group) B, exerts on B to improve the effectiveness of B's epistemic agency, either by putting B in the conditions to acquire an epistemic good or by preventing B from developing various forms of epistemic deficiencies (e.g., cognitive biases, unjustified beliefs, or inappropriate heuristic reasoning). Second, epistemic paternalism involves A's *interposition* with B's agency to the extent that B lacks the opportunity to exercise their own judgment in the way they think to be most appropriate.

Both components of Goldman's view of epistemic paternalism are featured in Ahlstrom-Vij's recent account (2013), according to which A undertakes an epistemically paternalistic practice towards B by doing (or omitting to do) X if and only if the following conditions are met:

(a) Doing X interferes with the epistemic autonomy or freedom of B to conduct inquiry in whatever way they see fit (*interference condition*);
(b) A does so without consulting B on whether B should be interfered with in the relevant manner (*non-consultation condition*); and
(c) A does so for the purpose of making B epistemically better off (*improvement condition*).

Let us consider in detail these three necessary and jointly sufficient conditions for an interference to be epistemically paternalistic. The interference condition captures Goldman's point on A's interposition with B's agency. Assuming an involuntaristic framework, according to which we cannot believe things on command, interfering with one's epistemic agency,

particularly with one's inquiry,[67] amounts to compromising one's freedom to choose the most appropriate methods and strategies to perform some epistemic task and thus to attain an epistemic good. The jurors case mentioned above amounts to an example of an external constraint on information access in which jurors' freedom to evaluate the case is compromised by the fact that the judge withholds relevant information in order to prevent them from becoming biased against the defendant. The prediction-model case features a constraint on information collection because it forces clinicians to collect some, and not other, information about patients and to ground their diagnosis on the results provided by the model. The mother case is an example of a slightly different sort of interference, in that her intervention affects her son's freedom to decide how to deal with his academic and professional interests by making things easier for him to opt for studying.[68]

According to the non-consultation condition, for an interference to be (epistemically) paternalistic, A does not ask B whether B is happy with A's interference. On Ahlstrom-Vij's account, for the interference to be paternalistic it is not necessary that B would object to A's interference, had they been consulted, nor that B would not welcome the interference itself. What matters instead is that A acts irrespectively of what B might think about the interference—that is, that A does not ask for B's opinion, or, in case A knows it, disregards what B wants.[69]

The improvement condition captures Goldman's point on protection, yet it goes beyond that concept, for it explains that the scope of A's interference is not merely that of protecting B from a potential epistemic harm. Rather, it aims at ensuring that B's epistemic agency benefits from the interference. One plausible way to account for how a paternalistic interference can make one epistemically better off is to refer to the notion of epistemic value as conceived by Pritchard (2009, 2013b), amongst others. His take on epistemic paternalism sheds light on a relevant weakness of Goldman's and Ahlstrom-Vij's perspectives, which measure the epistemic benefits and harms of a paternalistic interference in a purely veritistic way—that is, by considering the number of true beliefs that A allows B to acquire, or the number of false beliefs that A prevents B from acquiring (see 2013: 4). On a broader perspective, it sounds reasonable to concede that A's interference can make B epistemically better off in at least two more ways. First, B might improve their understanding of some subject matter *x* or avoid worsening their

[67] On Ahlstrom-Vij's view, inquiry cannot be reduced to belief formation; rather, it is something the subject does and whose purposes, methods, and activities "are selected specifically on account of their epistemic merits, that is, because of how they (as far as we can tell) tend to lead us towards true belief and away from false beliefs" (2013: 40).

[68] For further considerations on different kinds of constraints, see Bullock (2018: 2-3).

[69] Ryan (2016) makes a similar point about general paternalism when he suggests that Dworkin's condition (2010)—according to which A's interference is paternalistic insofar as A acts without the consent of B—be replaced with the requirement that A acts irrespective of the consent of B.

understanding of *x* as a result of A's interference. Second, B might acquire intellectual virtues or avoid forming epistemic vices because of A's interference. Notice that the improvement condition constrains A's purpose, rather than the outcome of A's interference. Thus, as Ahlstrom-Vij points out, for an interference to count as epistemically paternalistic it is not necessary that A promote B's wellbeing: A's failure in improving B's epistemic wellbeing might make the interference unjustified, but it does not affect its status as an epistemically paternalistic practice (2013: 49).

If we take a quick look at the debate on general paternalism, it is easy to notice that all three requirements for epistemic paternalism might be questioned. Ryan would presumably reject the non-interference condition, as he contends that "an action may be paternalistic without interfering in the liberty or autonomy of the object of the paternalist action" (2016: 126). Feinberg might reject the non-consultation condition since he believes that a coercive rule legislated for someone's sake and approved by the subjects interfered with is not paternalistic (1989: 20). And Dworkin (2017) would replace the improvement condition with a success-based condition, according to which A's interference shall improve B's welfare or promote B's interest, values, or good. In what follows, I assume that Ahlstrom-Vij's conditions can nonetheless be defended; I shall leave a more detailed discussion of these requirements for another time.

All I have said so far concerns the requirements for one's interference with another's agency to count as epistemically paternalistic. However, for an epistemically paternalistic interference to be justified, some story has to be told about how the interference comes to have the relevant beneficial effects it is meant to generate. That is a very complicated matter because it has to be shown not only that (i) an epistemically paternalistic interference is likely to promote the interfered-with subjects' epistemic good, but also that (ii) it does not damage their overall welfare. I shall explore this issue in detail in the next chapter but, for the sake of clarity, let me anticipate some relevant aspects here. The former requirement amounts to demonstrating that A must have a justified belief that their interference is likely to be beneficial for B. To account for this requirement, Ahlstrom-Vij introduces the *burden-of-proof condition*, which demands that "the would-be interferers are able to make a case that available evidence suggests that it is *highly likely* that everyone does or will benefit from the relevant form of interference, compared to relevant alternatives" (2013: 122). The latter requirement amounts to showing that the epistemic reasons for interfering do not clash with other relevant epistemic or non-epistemic reasons against intervention.

For the sake of argument, let us grant that (i) is fairly unproblematic, and focus on (ii). The first problem with this requirement is that any interference seems to violate at least the interfered-with subject's own autonomy or personal sovereignty. Some would presumably argue that such a violation might constitute a sufficient reason not to undertake any form of epistemically paternalistic interference. Yet there are ways to resist this objection. According to Bullock, the proponent of epistemic paternalism might respond that personal sovereignty is only pro tanto valuable, as there might well be circumstances in which our reasons for interfering outweigh the concern for one's autonomy (2018: 443). According to Pritchard (2013b), epistemic paternalism need not clash with this legitimate concern, because a small violation of someone's autonomy today might be justified by the fact that it leads to improving their freedom and autonomy in the longer term.

The second problem with (ii) pertains to finding a compelling way to cash out this requirement. Ahlstrom-Vij's strategy amounts to introducing the *alignment condition*, according to which in order for an epistemically paternalistic interference to be justified, A's epistemic reasons for the interference need to be aligned with A's non-epistemic reasons for the interference by either constituting additional reasons for interfering or by being silent on the issue—that is, by not constituting reasons against interfering (2013: 117). This condition does not require—as rival options do[70]—that A knows the weight of the reasons to be balanced, but only their valence (their direction for or against a given interference). However, it presents a relevant weakness: as Ahlstrom-Vij admits, it does not constitute a stable necessary condition for justified epistemic paternalism, because there may be cases in which a weak non-epistemic reason against interfering fails to outweigh robust epistemic reasons for intervention.[71]

Despite these problems, Ahlstrom-Vij holds that the burden-of-proof condition and the alignment condition are jointly sufficient to justify epistemically paternalistic practices (114). It is important to notice that his justification of epistemic paternalism targets large-scale situations in which someone's interference is going to have an impact on a considerable number of subjects. This explains why in his view the notion of evidence on the likelihood of an interference's beneficial effects is to be conceived in terms of statistical probability. Nonetheless, it seems possible to justify paternalistic interferences even in the absence of such statistical evidence, or so I shall contend. In several ordinary circumstances, some interferences can help particular subjects (or groups) in virtue of their specific epistemic situation. Think, for example, of a parent hiding a joke history book, i.e., a book including unreliable information

[70] See, e.g., Bullock's balancing-goods condition (2018: 440-441).
[71] See also Bullock (2018: 441-442) on this point.

and jokes about historical events, from their kids (Pritchard 2013b: 15); of a doctor breaking her patient's right not to know about his illness (Bullock 2018: 434), as she justifiedly believes he will benefit from knowing that he is out of danger; or of a teacher refraining from providing a student with the tools to solve a geometry problem in order to let her develop analytical skills. What matters in circumstances like these is that the interferer be an epistemically competent subject who has the ability to form justified beliefs about the benefits of the relevant form of interference as well as about why that course of action is meant to be more beneficial than relevant alternatives.

In this picture, Ahlstrom-Vij's burden-of-proof condition becomes a special instance of a more general requirement—call it the *epistemic-reasons condition*—according to which interferers must have robust epistemic reasons for believing that the subject(s) interfered with will benefit from the intervention, compared to relevant alternatives. Those who fulfill the burden-of-proof condition are by definition satisfying the epistemic-reasons condition too, yet in several cases someone can fulfill the latter without being in a position to satisfy the former. In such circumstances, an interference is justified insofar as it fulfills the epistemic-reasons condition and the alignment condition. I shall argue that for a paternalist interferer to be justified in intervening in someone else's inquiry as required by these conditions, they need to be *virtuous* interferers (see §5.5).

5.3 Justifying Epistemic Paternalism: Against the Expertise Strategy

In several passages of his work on epistemic paternalism, Goldman suggests that expertise is a fundamental component for defending epistemic paternalism and argues that "to justify any particular instance of such paternalism … we must have grounds for taking the agent to be an expert" (1991: 128). This claim seems to suggest that in his view, expertise should be considered a necessary condition for one to be entitled to undertake epistemically paternalistic interferences towards another. However, if someone thinks that, by definition, experts fulfill the justification conditions of paternalistic interferences, cognitive expertise would become a sufficient requirement for justified epistemic paternalism. In this section, I shall argue that his notion of expertise is neither necessary nor sufficient for justifying epistemically paternalistic interferences.

In particular, I shall focus on the account of expertise that Goldman endorses in his 1991 article on epistemic paternalism, which can be summarized as follows:

EXPERT. A subject S is an expert in a domain D iff
[T-LC*] S has true answers to core questions in D; or
[AC] S has the capacity to acquire true answers to core questions arising in D.

The structure of this account resembles the research-oriented one I have extensively analyzed in chapter 2, but it is worth pointing out a few relevant differences. First, the veritistic flavor of Goldman's perspective is ensured by a *truth-linked condition* that, unlike cognate possession conditions on expertise previously analyzed, takes into consideration the relevance of the questions that an expert is able to address in their domain of expertise and fails to make explicit reference to the comparative dimension of expertise.

Second, the definition includes an *ability condition* that allows us to also consider experts those who have the ability to solve new problems arising within D, no matter whether they actually have done so. The reader might wonder why this condition is proposed as part of what Goldman himself would call the possession condition on expertise instead of as a separate functional definition of an expert. The answer to this legitimate question is straightforward: in his 1991 work on expertise, Goldman had not endorsed the functionalist approach which we can find in more recent papers of his (see, e.g., 2018) and therefore he considered the truth-linked condition and the ability condition on a par qua requirements for the possession of expertise.

Third, and finally, Goldman has it that for one to be an expert it is enough that one fulfills either of these two conditions. Here, a difference in the formulation of the relationship between possessing true beliefs in one's domain of expertise and having the ability to acquire true answers to the questions of the domain need not reflect a change in Goldman's conception of such relationship. As a matter of fact, in his most recent accounts of an expert, it is still the case that fulfilling the truth-linked condition—albeit one that includes a comparative dimension of expertise—enables one to address new questions arising in a given domain.

That said, it would not be hard to recast this account of an expert along the functionalist approach that Goldman himself endorses in recent work. On a functionalist perspective, T-LC* constitutes the possession condition that specifies what it takes for an epistemic subject to fulfill the research-oriented function introduced in §2.2:

[RO-CAP] S has the capacity to contribute to the epistemic progress of D. S can provide such help by offering true answers to the questions under dispute in D.

Now, let us go back to the conditions for justified epistemic paternalism in light of this account of an expert. Someone, if not Goldman himself, might want to hold the view that experts satisfy the burden-of-proof condition. It could be argued that a subject A, in virtue of their expertise, is the best candidate to evaluate available evidence in D and make a case that their interference with B is highly likely to make B epistemically better off. In fact, it is plausible to contend that experts know much better than laypeople how to assess pros and cons of a given course of action in their domain of expertise in light of relevant alternative practices. Similarly, one might hold that experts are also better placed to evaluate potential non-epistemic reasons against intervention, thereby being in a position to satisfy the alignment condition.

On careful analysis, though, it is far from clear that Goldman's account of an expert can accommodate the requirements for justified epistemic paternalism. Let us focus on whether one's expertise is *sufficient* for them to be justified in undertaking epistemically paternalistic interferences. Consider the following example:

PROF. EVERYT SOLVED. Suppose Joseph is a young mathematician based at MIT who is working at Hilbert's problem n.3. Joseph knows that this specific problem has already been solved and knows its solution. His supervisor, Prof. Everyt Solved, is not only one of the most important mathematicians who worked at Hilbert's problems but is also well known for her distraction and insensitivity to others' epistemic needs. During a meeting with Joseph, Everyt suggests that Joseph try to work at problem n.3 from the beginning, as if it were still unsolved, in order to understand fully its structure and solution. A week later, she stops by Joseph's desk and, once she notices he is working on the problem, she tells him straight away the solution to the problem without allowing him to say a word.

I argue that this case satisfies the three conditions for an interference to count as epistemically paternalistic. Everyt's intervention fulfills the interference condition in that she limits Joseph's autonomy to conduct inquiry into Hilbert's problem n.3 in whatever way he sees fit. It also fulfills the non-consultation condition because she neither takes into consideration Joseph's opinion regarding her intervention nor consults him on the issue. Finally, it fulfills the improvement condition: due to her distraction and insensitivity to Joseph's needs, Everyt forgot

the advice she gave to Joseph and now she genuinely interferes to help him by offering the solution to the problem. I also maintain that Everyt satisfies Goldman's account of an expert: she not only has lots of true answers to the core questions in mathematics, but she also possesses the ability to contribute to the epistemic progress of the discipline, as her outstanding list of recent and forthcoming publications shows.

Nonetheless, Everyt's intervention does not constitute a justified case of epistemically paternalistic interference. Her complete insensitivity to the student's epistemic needs prevents her from being in a position to satisfy the epistemic-reasons condition. As a matter of fact, she might possess epistemic reasons for believing that Joseph will benefit from her intervention and that any further non-epistemic reasons for intervening align with the epistemic ones. Nonetheless, the problem lies with the relevant-alternatives component, as she definitely fails to evaluate which attitude is going to help Joseph get the most out of his intellectual inquiry. A more careful evaluation would have easily allowed Everyt to acknowledge that the strategy of letting Joseph work at Hilbert's problem on his own would have made him improve his understanding of the solution in a way that Everyt's intervention obviously cannot. Yet, assessing what course of action is going to be more beneficial for the subject interfered with is by no means a condition for one to be an expert. Thus, the example shows that someone's expertise does not ensure that their epistemically paternalistic interferences are justified.

Now, let us analyze whether being an expert is nonetheless a *necessary* condition for justified epistemic paternalism. The fundamental problem with the above definition of an expert is that neither of its conditions can ensure that an expert A has good reasons to think A's interference will make B epistemically better off, for they pertain to someone's having extensive knowledge in a given domain. Instead, what is required for an epistemically paternalistic interference to be justified is that the interferer have a clear view on what is epistemically better to do on behalf of B. Consider the following example:

> VIRTUOUS COLLEAGUE. Suppose Emma and Frank are in charge of the recruiting process for a big company that is hiring twenty new employees. Having recently noticed that Frank has developed a bias against female applicants, this time Emma wants to help him. So she asks for Frank's help with the first step in the selection process: evaluating the CVs of two hundred applicants and selecting the best forty profiles. But she provides him with blind CVs in order to prevent his bias from affecting his judgment, and she reveals the identity of the applicants only after he completes his task.

Again, this case satisfies the three conditions of epistemic paternalism: it fulfills the interference condition because Emma prevents Frank from conducting inquiry in whatever way he sees fit; it fulfills the non-consultation condition since Emma does not consult Frank on the issue of whether he would be happy if she blinded the applicants' CVs for him; and it fulfills the improvement condition, as Emma intervenes with the aim of improving his epistemic agency, namely of enhancing the chances that he get to know which are the most suitable profiles for the job positions that he needs to fill in.[72] Furthermore, one might want to argue that her interference is justified, for the available evidence suggests it is highly likely Frank will benefit from the interference compared to relevant alternatives—such as letting him select candidates on the basis of his gender bias—and there are no relevant reasons against intervening that Emma should take into consideration. Nonetheless, it seems clear that the justification of Emma's interference does not depend on any specific kind of expertise she might have. Rather, what puts her in the best position to fulfill her function as a paternalist interferer is the fact that she is sensitive to her colleague's biased attitude and acquires good evidence of what is his best epistemic interest.

To further stress this point, let us consider which intellectual abilities allow one to fulfill their function in the epistemic community. I contend that the intellectual virtues required for one to fulfill Goldman's account of the expert are largely different from those that allow one to satisfy the justification conditions of an epistemically paternalistic interference. Experts need to possess what in previous chapters I have called *research-oriented abilities*: virtues that allow one to exploit their fund of knowledge to find and face new problems arising in their field of expertise, such as intellectual curiosity, intellectual creativity, open-mindedness, intellectual courage, firmness, and autonomy. On the other hand, a paternalist interferer is virtuous insofar as they possess *novice-oriented abilities*: virtues that allow them to properly address B's epistemic dependency on them, thereby putting them in a suitable position for knowing what is epistemically best to do in the service of B. This set of abilities includes traits such as sensitivity to B's epistemic needs, intellectual generosity, intellectual empathy, sensitivity to B's epistemic resources, practical wisdom, and maieutic ability.

This distinction bolsters the thesis that cognitive expertise is neither a sufficient requirement for one to be justified in undertaking epistemically paternalistic interferences nor a necessary one. Being an expert is not sufficient for justifying epistemic paternalism, because

[72] Clearly, Emma's interference would presumably have clear practical benefits, in that it could lead him to select better candidates. That is not in contrast with the epistemic benefit that her intervention provides, nor does it undermine the epistemic purpose of her actions.

a person's expertise does not ensure that they satisfy the epistemic-reasons condition, as they may well lack the ability to evaluate what among several options is the most epistemically beneficial way to interfere with someone's inquiry. Furthermore, being an expert is not even a necessary condition for justified epistemic paternalism, for as the case of the virtuous colleague shows, someone can fulfill the requirements for justified epistemic paternalism without being an expert.

5.4 Justifying Epistemic Paternalism: The Epistemic Authority Strategy

A plausible alternative to the idea that experts constitute the ideal profile of paternalist interferers is offered by the notion of an *epistemic authority* that I have defended in previous chapters. In the Razian conception of practical authority, the main function of authorities is to serve the governed (or the novice)—that is, to do something in their service (see Raz 1986: 56). On my pluralistic approach, an epistemic authority can fulfill their function in various ways, ranging from imparting true beliefs to a layperson B (Zagzebski 2012) and helping B weigh available evidence (Jäger 2016; Lackey 2018), to imparting understanding to B and leading B to improve their understanding of some subject matter on their own.

Let me briefly recall the account of an epistemic authority that I have proposed in chapter 3. Sticking to Goldman's functionalist approach, it seems evident that epistemic authorities are supposed to fulfill a *novice-oriented function*, according to a subject A is an epistemic authority for a subject B in domain D iff

> [NO-CAP] A has the capacity to help others (especially laypersons) solve a variety of problems in D or execute an assortment of tasks in D which the latter would not be able to solve or execute on their own.

The possession condition of epistemic authority can be summarized as follows:

> EPISTEMIC AUTHORITY. A subject A is an epistemic authority for a subject B in domain D iff
>
> [EPC] A is better epistemically positioned than B is in D; and

[AC*] A possesses at least sensitivity to B's epistemic needs.[73]

A quick comparison between this notion and that of an expert should clarify why in principle epistemic authorities look like a more convincing exemplar of paternalist interferers than experts.

The first distinction concerns the functional definitions and illustrates that these categories of epistemic subjects have different roles in the epistemic community. The function of epistemic authorities is to help the interlocutor(s) achieve epistemic goals, as required by NO-CAP. Yet they have no commitment to fulfill RO-CAP, which instead explains that the service of experts amounts to making a contribution to the epistemic progress of their field. This distinction has bearing on the possession conditions as well, on which the second difference between the two notions focuses. Instead of Goldman's truth-linked condition, the definition of an epistemic authority includes the epistemic-position condition (EPC), which differs from the former in at least two relevant ways.

On the one hand, EPC does not require that A be epistemically superior to most people in a domain, but rather just to their interlocutor. Thus, one can be an epistemic authority to another about some subject matter without being an expert on that topic. For instance, I might be an epistemic authority for my mother on the history of Scotland simply because I have some vague knowledge of the main battles and events that happened there in the modern era while she knows nothing about this topic, but that would not make me an expert in that domain. Indeed, I could not contribute to the progress of the historical research in this field, as required by RO-CAP, yet I would still provide my mom with information she lacks and is interested in acquiring and therefore I would fulfill NO-CAP. On the other hand, EPC does not limit one's epistemic superiority to another to the number of true propositions or core answers one has in a given domain. A can, in fact, be better epistemically positioned than B also by having a better understanding of D, by being more intellectually virtuous than B, or simply by having access to more (or better) evidence. Finally, experts and epistemic authorities need to possess different intellectual virtues. The former should be able to find true answers to the questions arising in their field and therefore need to possess research-oriented abilities. The latter, instead, merely

[73] Careful readers will have noted that the proposed possession condition replaces the original conscientiousness-based requirement with a requirement featuring the notion of epistemic position. I have offered a motivation for such a move in §3.6. Furthermore, the proposed definition makes no reference to the notion of an ongoing relationship between the authoritative subject and the novice. This is motivated by a concern with simplicity and elegance rather than by a conceptual worry. As a matter of fact, in all the cases that I will discuss in the remainder of this chapter the authoritative subject shares an ongoing relationship with their interlocutor.

have to be sensitive to the interlocutor's epistemic needs, thereby displaying one of the most important novice-oriented abilities.

Let me stress once again that the proposed pluralistic account of epistemic authority presents two fundamental advantages over rival views:[74] first, it does not restrict the kind of *practice* that A is entitled to adopt towards B to either imparting information or advising in a more indirect way; second, it does not restrict the *epistemic good* provided by A to either true belief or understanding. Both features are crucial to our argument, in that, first, they allow us to infer that this account of epistemic authority could well include "undertaking paternalistic interferences" as one of the viable ways A can help B achieve epistemic goals in D. Second, they make room for the idea that paternalistic interferences can benefit the interfered subject in various ways. Finally, the differences between experts and authorities discussed in previous chapters highlight that epistemic authorities are good candidates as exemplars of paternalist interferers because they need to possess a fundamental intellectual virtue, namely the sensitivity to B's needs, that makes A care about discovering what is epistemically better to do in the service of B. For these reasons, we shall consider whether it might be the case that epistemically paternalistic interferences are justified insofar as the interferer is an epistemic authority.

5.5 Virtuous Paternalist Interferers

In this section, I will first introduce and discuss two objections to the plausible thesis that epistemic authorities are the ideal profile for paternalist interferers in the epistemic realm. I shall argue that, despite directing us on the right track, this thesis needs to be refined. Then, I will offer a more compelling account of paternalist interferers.

The first problem affecting this thesis sheds light on a relevant asymmetry between the functional definition of epistemic authorities and the functional definition of paternalist interferers. The asymmetry arises in cases where an epistemically paternalistic interference can be beneficial to some extent yet is not grounded in A's judgment of what is epistemically better to do in the service of B. Let us consider the following modified version of the mathematicians example introduced before:

[74] Compare it, in particular, with Lackey (2018) and Zagzebski (2012).

PROF. EVERYT SOLVED*. Suppose again that Joseph is working at Hilbert's problem n.3. This time, he knows that the problem has already been solved, but he does not know the solution. Thus, he decides to work at the problem as if it were still unsolved in order to achieve a deep understanding of its structure. Everyt, his supervisor, stops by Joseph's desk and, once she notices that he is working on that particular problem, reveals the solution without allowing him to say a word.

This example features a case in which Everyt clearly fulfills N-OF because she has the ability to help Joseph achieve an epistemic good that he lacks: knowledge of the problem's solution. She also displays some sort of sensitivity to the student's needs, as she provides him with some useful piece of information. Nevertheless, I contend that her interference is not justified, because Everyt did not consider the impact of her intervention on Joseph's inquiry nor has she formed any justified belief about alternative courses of action that she might have undertaken, as the epistemic-reasons condition requires. As a matter of fact, we expect a virtuous interferer at least to consider that letting Joseph work on the problem on his own would possibly allow him to achieve both knowledge of the solution and understanding of its structure at the same time with small risk, as he could look the solution up online and get to know how Max Dehn solved this problem in 1900 by appealing to invariants of polyhedra.

Thus, the fundamental asymmetry between the two functional definitions of epistemic authorities and paternalist interferers can be highlighted as follows. On the one hand, NO-CAP is neutral with respect to the epistemic goal Everyt should paternalistically help Joseph achieve. On the other, it is not sufficient that an interferer A has the ability to help B achieve *some* epistemic goods in D, nor that A has the ability to help B achieve *what B aims at* achieving. Paternalist interferers have a more specific service to fulfill: namely, they need to be able to help B achieve *what is epistemically better* for B in D.

The second objection challenges EPC by raising the doubt that paternalist interferers need not be better epistemically positioned than interfered subjects to successfully provide their service. Consider Emma's profile again in the case of the virtuous colleague. In the last section, I argued that her interference satisfies the conditions of epistemic paternalism introduced in §5.2 and that it could constitute a justified paternalistic interference. Nonetheless, some might argue that Emma is not an epistemic authority for Frank because she is not epistemically superior to Frank in any relevant sense with respect to the matter at issue. Indeed, both Frank and Emma have worked for decades in the HR department of that company, and already proved to be experienced recruiters. Then we should conclude that EPC is not a necessary condition

for one to be justified in undertaking epistemically paternalistic interferences, which would reinforce the thesis that epistemic authorities are not the ideal profile of paternalist interferers.

I am not interested in resisting this conclusion, as I have already argued that the thesis introduced at the beginning of this section needs some refinement. However, I shall undermine the claim that the notion of a paternalist interferer needs no EPC. On the broad conception of epistemic superiority endorsed here, one can be better epistemically positioned than another in a very local way—for example, by lacking a relevant bias or being able to spot it in other people in a given circumstance. In fact, the interferer, unlike the subject interfered with, displays the virtue(s) of epistemic justice, which enables the former to be sensitive to the latter's bias and triggers further intellectual virtues, such as the sensitivity to others' needs and epistemic resources, through which they figure out how to help the subject interfered with.

These objections shed light on the fact that paternalist interferers ought to satisfy more stringent requirements than epistemic authorities, both for what concerns the functional definition and for the intellectual virtues they need to display. As for the former, this section has provided support for the idea that a paternalist interferer needs to be able not merely to provide some kind of epistemic benefit for the subjects interfered with, but rather to figure out what is epistemically better for them to do in given circumstances. A full account of a paternalist interferer's function, though, also requires that the interferer be able to evaluate whether the epistemic benefits their interference is likely to generate are in line with further non-epistemic reasons against intervening that they might have, as required by the alignment condition. To clarify this point, consider the following revised version of Emma's case:

> VICIOUS COLLEAGUE. Suppose again that Emma and Frank are in charge of recruiting twenty new employees for a big company, and that Emma has recently noticed that Frank has developed a bias against female applicants. However, she also knows Frank is always suspicious of other colleagues and that her move will raise his worries and undermine his trust in her. Nonetheless, she provides him with 200 blind CVs, in order to prevent his bias from affecting his judgment, and asks him to select the best forty profiles.

In the example, Emma's interference cannot be justified, because it clashes with the available evidence for what is better to do on behalf of Frank in the given circumstances. Specifically, she lacks a necessary concern with maintaining a relationship of mutual trust among

colleagues—that is, a non-epistemic reason that does not align with the epistemic reasons supporting Emma's interference.

Both the function requirement and the definition requirement of a paternalist interferer should reflect Emma's lack of justification. I propose the following account of virtuous paternalist interferers, whose functional definition maintains that a subject A is a virtuous paternalist interferer for a subject B in domain D iff

> [NO-CAP*] A has the capacity to help B achieve what is epistemically better for B in D when permitted by the balance of reasons.

Thus, I put forth the following possession condition:

> VIRTUOUS PATERNALIST INTERFERER. A subject A is a virtuous paternalist interferer for a subject B in domain D iff
>
> [EPC] A is better epistemically positioned than B is in D;
> [VC] A's judgment about how to interfere with B's inquiry is the product of A's cognitive faculties; and
> [AC**] A deploys a wide range of novice-oriented abilities in judging how to intervene.

As should be evident, this definition differs from the account of epistemic authority in two respects. First, the definition, unlike that of an epistemic authority, includes a *virtue condition* (VC), which ensures that the interferer's decision to intervene arises out of a competent use of their cognitive faculties. This requirement is necessary to avoid parallel Gettier-style cases, where the fact that A's interference fulfills the epistemic-reasons condition and the alignment condition is simply a matter of luck. Imagine, for example, a case in which a doctor breaks a patient's right not to know the result of a medical test, because she has a justified belief that he will benefit from knowing that he is in good health. Unbeknownst to her, someone replaced the result of the test with someone else's. As it turns out, the two tests had identical results. Thus, the doctor's reasoning and judgment about whether to interfere with his agency would not have been different had she considered the correct results. Yet her beliefs about both the patient's health situation and the best way to intervene are not justified, as they are true simply because

of luck. A virtue condition allows us to ensure that a case like this does not become an example of justified epistemic paternalism.

Second, the definition of a virtuous paternalist advisor includes a stronger version of the ability condition, which is meant to accommodate the fact that paternalist interferers have a different, stricter function to fulfill than epistemic authorities. Indeed, for one to be in a position to accomplish NO-CAP* it is necessary that one not only be sensitive to another's epistemic needs, but also that one be able to weigh the epistemic benefit against possible non-epistemic reasons not to interfere—thereby displaying practical wisdom—and to make sure that the subject interfered with is in a position to take advantage of the epistemic good one tries to provide them with—thereby displaying sensitivity to their epistemic resources.

This account allows us to explain why Emma is a virtuous paternalist interferer in the first example (the virtuous colleague) while she fails to satisfy the requirements in the second case (the vicious colleague). Specifically, in the former scenario, Emma not only fulfills EPC, as we have seen early on in this section, but also VC and AC**, since her judgment about how to help Frank is the product of her cognitive faculties and she proves herself able to make a virtuous use of her novice-oriented abilities by determining that the balance of reasons favors her intervention. For this reason, she proves herself able to accomplish NO-CAP*, and therefore she can be considered a virtuous paternalist interferer. In the latter scenario, instead, Emma fails to fulfill both the functional definition and the possession conditions. As for the former, her inability to evaluate whether the epistemic reasons for interfering are aligned with the non-epistemic ones shows that she is unable to figure out what is epistemically better to do in Frank's service. As for the latter, although her judgment is the product of her cognitive faculties and she is better epistemically positioned than her colleague because she is aware of his bias against female applicants, she does not fulfill AC**. For the fact that Emma has low concern for the potential harms of her interference indicates a failure to exercise her novice-oriented abilities in a virtuous way—particularly, practical wisdom as well as sensitivity to his needs and resources.[75]

[75] A similar story could be told to explain why Everyt, in the case of Prof. Everyt Solved*, is not a virtuous paternalist advisor. Everyt does not fulfill NO-CAP*, because she fails to display the capacity to weigh the epistemic and the non-epistemic reasons for interfering with Joseph's inquiry. Even though she is better epistemically positioned than him in the domain of Hilbert's problems, she proves to be insensitive to his needs and resources, and she lacks the ability to guide him in the right direction without doing all the work on his behalf.

5.6 Benefits of the Account and the Challenge of Institutional Interferers

Before concluding this inquiry into the features of paternalist interferers in the epistemic realm, I want to shed light on several benefits of the proposed account and address one final challenge.

First, this account of a virtuous paternalist interferer settles the original worry with Goldman's thesis that one can undertake epistemically paternalistic interferences insofar as one is an expert. I showed that his account of an expert underestimates a fundamental feature of paternalist interferers, namely that they need to have justified beliefs on what is epistemically better for the subject interfered with to do in given circumstances. Thus, paternalist interferers, unlike experts, need to care about the epistemic wellbeing of the subject interfered with as well as to display and appropriately exercise novice-oriented abilities.

Second, the account goes beyond Goldman's and Ahlstrom-Vij's limited veritistic perspective to endorse a broader view of the epistemic wellbeing that includes both knowledge and understanding as valuable epistemic goods promoted by paternalistic interferences, as suggested by Pritchard (2013b), among others.

Third, this inquiry into who can justifiedly undertake epistemically paternalistic interferences accounts for a special way to offer guidance in the epistemic realm and therefore individuates a peculiar way in which an epistemically superior subject may do something in the service of another. This analysis shows that the function of paternalist interferers should not be reduced either to that of experts or to that of epistemic authorities, as the former, unlike the others, need to be able to balance epistemic and non-epistemic reasons for intervention on behalf of someone who may not be aware that another is interfering with their own agency.

Despite the benefits of the proposed account of paternalist advisors, one might worry that this view has a very limited scope in that it only applies to one-to-one relationships between two epistemic subjects, one of which is better epistemically positioned and more intellectually virtuous than the other. However, any plausible account of epistemic paternalism should take into consideration cases in which groups and institutions undertake epistemically paternalistic interferences towards one or many epistemic subjects. That might be the case, for example, with a state imposing compulsory school age and with a health department mandating the introduction of prediction models for medical diagnosis and prognosis in order to prevent clinicians from overestimating their expertise and clinical abilities. Setting aside whether these particular interferences may or may not be justified, one might contend that the account I endorse cannot explain who would be a virtuous paternalist interferer in similar cases.

The considerations on group epistemic authority and expertise that I have offered in chapter 4 suggest that this worry can be resisted and provide support for the thesis that the virtuous paternalist interferer can be extended to accommodate cases of group and institutional paternalist advisors. I contend that this can be the case because there are available ways to extend EPC and VC to groups, but more importantly because, as I have already shown, collective bodies can display intellectual virtues and, more specifically, novice-oriented abilities.

In light of the arguments offered in §4.2.1, it looks as though EPC can be applied to the case of collectives because a group or an institution can be better epistemically positioned than others in their activity as collectives and institutional structures. This is possible at least to the extent that some groups possess more knowledge or better understanding than others, as in a case in which a scientific team conducting experimental research on the benefits and the harms of eating red meat is better epistemically positioned on the topic of the impact of food on human health than a farmers association.

Collectives can also possess novice-oriented abilities, or so I have argued in §4.2.2 and §4.4, insofar as their members jointly commit to the motive or the end of a virtue and the collective as a whole proves to reliably achieve the end of such virtue. A few more words are necessary to explain how the virtue condition can be extended to collectives. I shall limit myself to mentioning that Kallestrup has recently defended a collective-virtue epistemology along the lines of Sosa's performance epistemology. More will have to be said on these topics, but what interests us here is the possibility of arguing that a group can have knowledge "when the truth of its belief is a product of its innermost competence … in suitable shape and situation" (2016: 10) provided that the group forms the appropriate joint intentions (12-13). From the combination of the basics of a collective-virtue epistemology and Kallestrup's remark that the epistemic aims of a group may well extend beyond the acquisition of knowledge (15), it follows that an institution can fulfill VC, if its judgment about whether to interfere with someone's inquiry for their own epistemic benefit manifests its innermost competence in suitable shape and situation. Thus, there are ways to prevent a Gettier-style case of an institution's epistemically paternalistic intervention into someone's inquiry from being justified.

In conclusion, these considerations show that the three conditions of the virtuous paternalist interferer can apply to groups and institutions, thereby reflecting the idea that even a collective can be justified in undertaking epistemically paternalistic interferences insofar as it displays (a) appropriate competence in judging whether to interfere, (b) some sort of epistemic superiority to the subject interfered with, and (c) those virtues that put it in a position to evaluate what is

epistemically better to do in their service. The project of extending this account of paternalist advisors in the epistemic domain to institutions should also take into consideration that the bigger is the number of the subjects affected by an institution's interference, the more complicated will weighing epistemic and non-epistemic reasons for intervention be. As a matter of fact, it seems reasonable that it takes more for an institution interfering with the agency of hundreds or thousands of people to fulfill the requirements of a virtuous paternalist advisor than it does for subjects like Everyt or Emma in our examples.

5.7 Conclusion

This chapter has explored the fairly new topic of epistemic paternalism from an original point of view, namely by analyzing the requirements that allow one to be considered a virtuous paternalist interferer in the epistemic realm. After introducing the conditions for an interference to be epistemically paternalistic and two prominent strategies for justifying epistemic paternalism (§5.2), I argued against Goldman's thesis that experts are the most appropriate candidate as paternalist advisors (§5.3). In §5.4, I suggested the notion of an epistemic authority—instead of that of an expert—as a plausible ground for an account of a paternalist interferer. As it turned out, the proposed view suffered from several problems, yet it allowed us to shed light on a fundamental feature that virtuous paternalist advisors need to exercise, namely the capacity to help subjects interfered with achieve what is epistemically better for them in a given situation. Based on this consideration, I proposed a virtue-based account of the paternalist interferer and showed how it accommodates common cases of epistemically paternalistic interferences (§5.5). Finally, I highlighted the benefits of the account and showed how it can apply to cases in which the interferer is a group or an institution (§5.6). My ultimate hope is that this chapter not only outlines an original answer to the Normative Question concerning the rational boundaries of the interaction between non-peer agents, but also that it contributes to showing how virtue epistemology and, in particular, the study of intellectual virtues can contribute to research in (social) epistemology as a whole.

Chapter 6

Epistemic Paternalism, Personal Sovereignty, and One's Own Good

Abstract. A recent paper by Bullock (2018) raises a dilemma for proponents of epistemic paternalism. If epistemic paternalists contend that epistemic improvements contribute to one's wellbeing, then their view conflates with general paternalism. Instead, if they appeal to the notion of a distinctive epistemic value, their view is unjustified, in that concerns about epistemic value fail to outweigh concerns about personal sovereignty. In this chapter, I address Bullock's challenge in a way that safeguards the legitimacy of epistemic paternalism, albeit restricting its scope to a limited range of cognitive projects. After shedding light on a problem with how Bullock singles out cases to which the dilemma applies, I argue that there is at least one reasonable way of interpreting the notion of "personal autonomy" which legitimates and justifies undertaking epistemically paternalistic interferences for one's epistemic good.

6.1 Introduction

One of the important issues about epistemic paternalism that remains unsolved has to do with motivating epistemic paternalism, that is, with providing an explanation of *why* epistemically paternalistic interferences are justified. Addressing this question involves giving an account of the value of the epistemic improvements that these interventions are meant to bring about. In her "Knowing and Not-knowing for Your Own Good: The Limits of Epistemic Paternalism", Emma Bullock raised a neat dilemma for epistemic paternalists, one that puts pressure on the very possibility of offering a motivation for paternalistic interferences in the domain of inquiry. In a nutshell, if epistemic paternalists contend that epistemic improvements contribute to one's wellbeing, then their view conflates with general paternalism. Instead, if they appeal to the notion of a distinctive epistemic value, their view is unjustified, in that concerns about epistemic value fail to outweigh concerns about personal sovereignty.

The goal of this chapter is to address Bullock's challenge in a way that safeguards the legitimacy of epistemic paternalism, albeit restricting its scope to a limited range of cognitive projects. This project is meant to complete the answer to the Normative Question that I have

already begun to address in chapters 3 and 5. After briefly recalling the definition of epistemic paternalism and reconstructing Bullock's dilemma (§6.2), I shall attempt to fulfill the chapter's aim by way of two main moves. First, I shall take issue with the setup of Bullock's dilemma and, in particular, with how Bullock singles out cases to which the dilemma applies. Although I agree with her that the dilemma should not apply to standard educational settings involving young children—since in most cases of this sort, epistemic paternalism is justified—I shall point out that her proposed criterion for excluding these cases is highly problematic (§6.3). Then, I shall offer a solution to the dilemma on behalf of those who find scope for justification of epistemic paternalism in the distinctive value of some interferences. Specifically, I shall argue that there is at least one reasonable way of interpreting the notion of "personal autonomy" which legitimates and justifies undertaking epistemically paternalistic interferences for one's epistemic good (§6.4).

6.2 Bullock's Dilemma

As I have discussed in the previous chapter, a paternalistic interferer A undertakes an epistemically paternalistic practice towards a subject interfered with B by doing (or omitting to do) X if and only if the following conditions are met:

(i) Doing X interferes with the epistemic autonomy or freedom of B to conduct inquiry in whatever way B sees fit (*interference condition*);
(ii) A does so without consulting B on whether B should be interfered with in the relevant manner (*non-consultation condition*);
(iii) A does so for the purpose of making B epistemically better off (*improvement condition*).

The main difficulty with any attempt to motivate epistemic paternalism concerns the general account of epistemic value from which paternalistic interferences inherit their justification and its relationship with the broader notion of one's own good or wellbeing. Bullock distinguishes two approaches to epistemic paternalism depending on how each determines what counts as a genuine benefit to the subject interfered with. *Eudaimonic* epistemic paternalism confers value to paternalistic interferences based on their contribution to B's *overall wellbeing*. *Strict* epistemic paternalism, in contrast, confers value to paternalistic interferences based on the

epistemic benefit they generate for B. The two horns of Bullock's dilemma are basically arguments for why both views fail to provide a compelling motivation for undertaking paternalistic interferences in the epistemic domain.

Before introducing her arguments, it should be pointed out that Bullock's main targets are cases in which the individual interfered with is the one who benefits from the intervention—i.e., cases of *direct* epistemic paternalism—and is mature enough to make decisions about their own inquiry—i.e., cases of *hard* epistemic paternalism (2018: 434-435).[76]

Let us first consider Bullock's argument against eudaimonic epistemic paternalism. On this view, paternalistic interferences can generate epistemic improvements that are constitutive of wellbeing or merely instrumentally valuable to wellbeing if they simply increase the likelihood that B will be better off in terms of their wellbeing (Bullock 2018: 437). Either way, Bullock argues, eudaimonic epistemic paternalists will have a hard time showing how their view distinguishes itself from general paternalism. For not only cases of general paternalism fulfill the interference condition and the non-consultation condition. If we read the improvement condition along the lines of eudaimonic epistemic paternalism, some general paternalistic interventions fulfill it too, insofar as generating epistemic improvements is just one of the ways in which an interference can improve B's wellbeing.

But why should epistemic paternalists worry about keeping a neat distinction between the eudaimonic account and general paternalism? As Bullock rightly notes, if epistemic paternalism collapses into general paternalism, it is no longer clear what role the improvement condition is supposed to play (438). For it may well be the case that some paternalistic interventions improve B's overall wellbeing at the cost of sacrificing B's epistemic welfare and that granting legitimacy to such interferences undermines the scope of epistemic paternalism.

To name just one of Bullock's examples, self-enhancement bias can contribute to one's wellbeing by favoring an overly positive conception of oneself and thereby diminishing the risks of depression (439). Eudaimonic epistemic paternalists seem committed to granting that there may be cases in which interferences that promote this cognitive bias are justified in virtue of their contribution to B's overall wellbeing, despite the fact that such bias reduces or obstructs B's self-knowledge.

Strict epistemic paternalism is well positioned to avoid this unwelcome consequence of the eudaimonic account, in that it assumes that epistemically paternalistic interferences fulfill the improvement condition to the extent that they bring about a genuine epistemic benefit for B.

[76] Bullock contrasts these notions with those of *indirect* epistemic paternalism, i.e., interferences with an individual that aim to benefit others, and *soft* epistemic paternalism, i.e., interferences with an individual who lacks the competence to make decisions about their own inquiry.

Thus, strict epistemic paternalists—at least, on the veritistic approach Bullock seems more concerned with—would not concede that promoting self-enhancement bias can constitute a legitimate paternalistic intervention in the epistemic domain. Nonetheless, strict epistemic paternalism has to face a challenge of its own, in that it has to explain whether and how epistemically paternalistic interferences can bring about epistemic benefits without damaging B's overall wellbeing. Bullock's contention is that strict epistemic paternalists fail to provide a compelling argument of this sort.

Available answers to the challenge for strict epistemic paternalism go in two different directions. One is offered by the *balancing goods condition*, according to which an epistemically paternalistic interference is justified insofar as A has no good reason to suppose that their intervention will make B all-things-considered worse off. In other words, this condition requires that A weighs the epistemic benefits of their intervention with its other possible non-epistemic effects and confers justification to it to the extent that the reasons for intervention outweigh the reasons against interfering.

This requirement provides two important advantages, in that it accommodates cases in which A's interference aims at making B better off both epistemically and in some other relevant respect, and it avoids the risk that A's interference be justified when it brings about an epistemic benefit for B at the cost of damaging other spheres of B's wellbeing.[77] However, as Ahlstrom-Vij points out, in several cases weighing reasons might become an extremely difficult task because it might not be at all obvious how to put on the same scale values that differ in quality: in particular, in several cases it might be difficult to assess whether the epistemic reasons in favor of interfering outweigh the non-epistemic reasons against doing it (2013: 116-117).

To avoid this problem, Ahlstrom-Vij suggests to replace the balancing goods condition with the *alignment condition*, as I have anticipated in §5.2. According to this requirement, an epistemically paternalistic interference is justified only insofar as A's epistemic reasons for the interference are aligned with A's non-epistemic reasons for the interference, either by being additional reasons for intervening or by being neutral—that is, by not constituting reasons against intervening (117). Unlike the former condition, the alignment condition merely requires that A only knows the valence of the reasons involved in the evaluation—namely, the direction for or against a given interference.

[77] Ahlstrom-Vij calls this form of interferences *mixed paternalism* (2013: 115). An example of mixed paternalism is a judge's decision not to reveal information about a defendant's past record of crimes both to prevent the jurors from developing a bias against the defendant (epistemic benefit) and to safeguard the welfare of the defendant themselves (non-epistemic benefit).

Yet, this condition is problematic too, as there might be cases in which a weak non-epistemic reason against intervening would commit us to deny that an interference in favor of which we have strong epistemic and non-epistemic reasons can be justified. For example, suppose the surgical procedure through which I can donate my bone marrow to a friend is potentially dangerous given my actual medical conditions, and I tell the doctor I want to do it and do not want to be informed about any immediate or long term side effect I could suffer from. The doctor initially consents to my request, but after acknowledging that the surgery went just fine and reflecting that my decision not to know will sit badly with my hypochondriac tendencies, she disregards my request and informs me about the positive results of the procedure.

In this case, the doctor has a weak non-epistemic reason against intervening, i.e., my request not to be informed, and strong epistemic and non-epistemic reasons in favor of intervening, both having to do with the positive effects of knowing how the surgery went. It seems reasonable to suppose that the doctor's epistemically paternalistic interference is justified as the benefits of intervening override the harm of disregarding my request. If so, then the alignment condition fails to provide a necessary condition on justified epistemic paternalism. Ahlstrom-Vij is ready to concede that there might be a very minor range of situations to which the alignment condition does not apply. However, Bullock has shown that this weakness is more pervasive than Ahlstrom-Vij thinks. Basically, no paternalistic interference would be justified on this view because all interventions of this kind would have to outweigh one clear reason against interfering, namely the violation of B's personal sovereignty. For it should be clear that limiting or disrespecting one's personal sovereignty is a form of damaging one's wellbeing (2018: 441-442).

According to Bullock, this problem undermines the alignment condition, for neither of the replies available to the strict epistemic paternalist meets the challenge. On the one hand, strict epistemic paternalists could deny that we have to respect personal sovereignty, but this is an extreme view that would need to be supported by a strong argument. On the other, they could contend that personal sovereignty is only *pro tanto* valuable but, in that case, only weighing reasons would allow us to determine whether in a given situation the concern with personal sovereignty is outweighed by other available reasons for interfering with one's inquiry.

Since the alignment condition cannot be rescued from its weaknesses, it looks as though strict epistemic paternalists will have to address the concern with personal sovereignty with the resources provided by the original balancing goods condition. In this scenario, Bullock argues that they could either take an *extreme* route and contend that bringing about an epistemic benefit

for B is the only good that can outweigh personal sovereignty or opt for a *moderate* position. In particular, strict epistemic paternalists could argue that (a) the concern with personal sovereignty can be addressed by interventions that promote B's epistemic benefit *or* B's wellbeing, or (b) the concern with personal sovereignty can be addressed by interventions that promote *both* B's epistemic benefit *and* B's wellbeing (2018: 442).

All these views have their own problems. On the extreme position, it is far from clear why epistemic value should be the only kind of value that can outweigh personal sovereignty. Proponents of moderate option (a) shall show how the promotion of an epistemic benefit can trump personal sovereignty and the promotion of one's wellbeing when this contrasts with epistemic value. Finally, proponents of moderate option (b) shall show why it is necessary that an interference brings about both epistemic and non-epistemic value in order for epistemic paternalism to be justified.

The challenges for proponents of the moderate options, as Bullock admits, can in principle be met and therefore constitute the only available opportunity for epistemic paternalists—in fact, for strict epistemic paternalists—to solve the dilemma and justify epistemic paternalism as a legitimate practice (444). The remainder of this chapter will be devoted to outlining an argument in favor of this position.

6.3 Soft Epistemic Paternalism and the Case of Education

Before addressing the dilemma and outlining a potential way out for strict epistemic paternalists, I want to shed light on a general, yet related, issue that seems to be receiving less attention than it deserves in the current discussion on epistemic paternalism. As I pointed out in the last section, Bullock's dilemma takes stock of hard epistemic paternalism, that is, interventions aimed at bringing about some sort of epistemic benefit for a subject who is judged "to be worthy of having their decisions about their inquiry respected" (2018: 436). Bullock intentionally sets aside cases of soft epistemic paternalism, that is, circumstances involving individuals who lack the competence to make decisions about their own inquiry, and offers the case of a child's education as a paradigmatic example for this type of paternalistic interventions.

As I shall argue below, in principle there is nothing wrong with this restriction in scope. For as we have already made clear, this phenomenon is in itself a violation of one's autonomy and therefore a *prima facie* problematic intervention. Thus, we should hope that the epistemically paternalistic interferences that take place in a given environment are meant to

bring about an epistemic benefit only for individuals who lack the ability to make decisions about their own inquiry.

However, a closer look at the requirements of soft epistemic paternalism reveals that Bullock's criterion for distinguishing between hard and soft instances of epistemic paternalism looks more problematic than it initially appeared. The main problem with this criterion lies in the very notion of a competence to make decisions about one's own inquiry. Bullock takes education as a paradigmatic domain to clarify the distinction between the two kinds of epistemic paternalism, but education does not seem to fit the criterion. If it is clear that the child lacks the competence to make decisions about their own inquiry, it is not clear what it takes for a mature epistemic subject to have such competence.

Consider again the case of a parent moving a joke history book out of their child's reach to prevent them from acquiring inaccurate historical information (Pritchard 2013b: 15), or a case in which a teacher refrains from providing a student with the answer to a geometry problem in order to let them work out a solution and develop analytical skills. It seems plausible to contend that both situations constitute cases of soft epistemic paternalism, in that the subjects interfered with lack the competence to make decisions about their own inquiry and the educators' paternalistic interventions bring about an epistemic benefit for them.

Now consider another pair of cases involving mature epistemic subjects. The first scenario is a more fortunate version of PROF. EVERYT SOLVED (see §5.3), namely one on which the math professor does not reveal to the student that the problem he is working on is one of Hilbert's solved problems, nor does she provide him with the solution because she judges that letting the student work on the problem on his own would allow him to develop a better understanding of the problem and improve their mathematical abilities. The second scenario is VIRTUOUS COLLEAGUE, that is, the case in which Kate knows that her colleague Jerry—who is working with her at recruiting new employees for the company—has developed a bias against female candidates and decides to provide him with blind CVs in order to prevent his bias from affecting his judgment. Should we consider these examples as cases of hard epistemic paternalism? That is: are the student and Jerry mature epistemic subjects who have the competence to make decisions about their own inquiry?

At first glance, it looks as though we should reply in the affirmative: the subjects interfered with are neither children nor individuals with limited awareness of their autonomy in both the practical and the intellectual sphere. However, their competence to make decisions in the respective domains is severely limited. Given the early stage of his career, presumably the math student does not know whether it would be epistemically better for him to get a straightforward

solution to Hilbert's problem or to approach it as an extant problem in need of a solution. For he lacks the competence to evaluate what he could learn by dealing with the problem, what skills he would develop, and how this effort would reward him later in his career.

Similarly—in fact, even more dramatically—Jerry is completely blind to his cognitive bias. In this case, we should be wary of granting him the competence to decide about his own inquiry for two main reasons. First, because he lacks awareness of something relevant to possessing such competence. Second, because—as in the nature of biases in general—he would be reluctant to admit that he has developed one and/or is unable to change his epistemic conduct, were Kate to make him aware of his attitude towards female applicants.

Yet, if the characters of our toy examples lack the ability to choose what is epistemically better for them to do in the respective contexts, then it becomes at least dubious whether we should treat these cases differently than the ones involving children. As this argument illustrates, the boundaries between soft and hard epistemic paternalism become less and less neat. Hence, we are entitled to wonder what consequences this has on the kinds of epistemically paternalistic interferences that should be immune to Bullock's dilemma.

A tempting answer would be to extend the domain of soft epistemic paternalism so as to encompass many educational cases involving mature-yet-incompetent epistemic subjects. This move should not worry us too much, in that the overarching aim of education is no doubt epistemic and its main implications are typically confined to the intellectual domain.[78] The bad news for Bullock is that, as the cognitive bias example shows, the problem of separating soft and hard cases of epistemic paternalism can be pressing even outside education, where one's decisions may also affect the lives of others.

Thus, on the one hand, it seems as though there is scope for restricting the set of interferences that fall within soft epistemic paternalism as much as possible. On the other hand, Bullock's criterion for distinguishing between soft and hard cases of epistemic paternalism authorizes us to include a larger range of interferences in the set of soft cases, thereby exempting them from falling prey to Bullock's own dilemma. I must confess that I have no ready-made answer to this problem, which, to my knowledge, has been surprisingly overlooked to date. This problem, though, becomes somewhat relevant in relation to my proposed solution to Bullock's dilemma. So, I shall postpone any further comment on it until §6.5 and focus my attention on the discussion of Bullock's dilemma.

[78] It remains a highly disputed question in the epistemology of education what constitutes the primary epistemic aim of education (see, e.g., Baehr 2019; Kotzee, Carter, and Siegel 2019; Siegel 2018). What matters for the purposes of this chapter is simply to note that, despite ongoing divergences, epistemologists of education agree that the primary or overarching aim of education is an epistemic one.

6.4 Epistemic Paternalism Justified: On Personal Autonomy as Sovereignty

The solution I shall propose to escape the dilemma raised by Bullock constitutes an argument on behalf of strict epistemic paternalists embracing moderate option (a), as discussed in §6.2. More precisely, I will argue that in some cases epistemic paternalism can be justified because the promotion of an epistemic benefit trumps the concern with personal sovereignty both when such epistemic benefit aligns with one's overall wellbeing and when it clashes with one's wellbeing.

Let me be clear from the start: I am not going to sell the balancing goods condition as an easy requirement to fulfill here, as I have already clarified that it is far from being so. My goal is rather to single out a set of cases of (hard) epistemic paternalism whose structure is such that they do not fall prey to Bullock's dilemma if one is willing to endorse a broader account of personal autonomy than mere personal sovereignty. This result might look modest to a demanding reader who seeks a wide spectrum treatment for the issue at stake. Yet, it will provide us with enough concrete instances of justified epistemic paternalism to manage and, as I argued in §6.3, we might well have good reasons to keep this list relatively short.

The *personal sovereignty model* of personal autonomy revolves around "the idea of having a domain or territory in which the self is sovereign" (Feinberg 1989: 52). The boundaries of this territory presumably include any self-regarding decision, that is, any decision that directly affects only the interest of the decision maker (56). Any violation of the boundaries by uninvited individuals is illegitimate. If we take this account by the book, the right of self-determination is as morally basic as the good of self-fulfillment. What is more, self-determination may lead one to harm oneself and yet nobody else has a right to interfere with one's decisions, in that "autonomy is even more important than personal wellbeing" (59).

Clearly, this model of personal autonomy allows no room for any sort of paternalistic interference, be it general or epistemic. It might well be the case that Jerry's epistemic agency, in our example, would benefit from having personal information about the applicants removed from their CVs, but his right of self-determination outweighs Kate's concerns with his epistemic welfare, despite the fact that it will lead him to provide a biased assessment of the CVs.[79] Thus, if this is how we conceive of personal autonomy, (epistemic) paternalism will never be justified.

[79] Notice, though, that Jerry's right of self-determination does not outweigh the candidates' right to receive a fair assessment of their suitability for the job. Thus, although it might not be legitimate for Kate to interfere with Jerry's

However, other accounts of personal autonomy provide us with a different diagnosis of the relationship between self-determination and personal good, namely one in which personal sovereignty is only *pro tanto* valuable and, as a consequence, epistemically paternalistic interferences may be justified on specific grounds. Let us consider the *model of autonomy as a condition* outlined by Feinberg (1989), according to which autonomy is a *condition* that involves both the ability to make rational choices and *de facto* self-government, that is, the opportunity to exercise one's rights and capacities (31). According to this approach, autonomy requires several virtues: among them, Feinberg mentions moral authenticity, moral independence, integrity, self-discipline, self-reliance, but also those intellectual virtues that make one responsible for the self, namely intellectual courage, trustworthiness, reliability, and good judgment (44).

If we assess Jerry's example in light of this account, it is no longer clear that Jerry is such an autonomous individual. His overwhelming proneness to cognitive biases might undermine his ability to make rational choices and it certainly prevents him from displaying relevant virtues such as reliability, good judgment, and self-discipline—to the extent that being prone to his cognitive biases makes him "governed from the outside" (40).

However, if it is true that Kate's interference brings about an epistemic benefit for Jerry, our critic might still object that such benefit fails to counterbalance the violation of Jerry's right to self-determination (or personal sovereignty), which is necessary to provide strict epistemic paternalists with a way out of Bullock's dilemma. While this diagnosis does not suffice to settle the problem on its own, it points us in the right direction: as I shall argue in the remainder of this section, epistemic benefits trump concerns about personal sovereignty insofar as they contribute to one's overall personal autonomy, no matter whether on balance they also improve one's overall wellbeing.

Two brief remarks are in order here: first, following the model of autonomy as a condition, I assume that having autonomy is a matter of degree; second, I shall contend that becoming (more) autonomous may sometimes happen at the expense of one's wellbeing. Let me clarify these points while discussing another example involving a specific kind of social epistemic structure: *epistemic bubbles*.[80] Epistemic bubbles are structures of exclusion that prevent large groups of epistemic subjects from taking into due account some kinds of information. They impede distribution of a complete range of information by omitting relevant testimony from sources endorsing a rival perspective and therefore they distort the informational environment

agency for his own epistemic good, it might well be legitimate for her to intervene in order to protect the candidates' rights.

[80] The following characterization of epistemic bubbles is taken from Nguyen (2018).

of their members. Bursting epistemic bubbles requires that some members are exposed to excluded information.

Epistemic bubbles are particularly interesting in the discussion of epistemic paternalism because they not only degrade the epistemic welfare of their members, but they are also likely to put members' personal autonomy at risk. Consider the following example:

> THE CIA BUBBLE. Suppose Sarah has been raised in a small rural community that believes in all sorts of conspiracy theories regarding the corruption of the US government and the involvement of the CIA in events such as 9/11, the Kennedy assassination, and Malala Yousafzai's attempted murder. Sarah has been convinced that these conspiracy theories are reliable by made-up data and evidence as well as by acknowledging that all members of her community endorse these theories. After living several years in a small village, Sarah moves to the city to work and befriends Mary, who does not seem to have a problem avoiding any conversation about the U.S. government and the CIA, as Sarah explicitly requests. One day, Mary overhears a phone call between Sarah and her family and figures out that Sarah belongs to a community with the aforementioned features. Mary decides to intervene by inviting Sarah for dinner and presenting her with pieces of counterevidence for her conspiracy theories as well as with journal articles detailing the mechanisms typical of epistemic bubbles and their likelihood to proliferate in communities with such and such features.

Sarah's community constitutes an example of an epistemic bubble, in that their members have not been exposed to any sort of counterevidence and foster their belief in these conspiracy theories by discussing these issues solely with insiders. Furthermore, no matter the outcome of Mary's interference, this story exemplifies a case of epistemic paternalism: Mary infringes Sarah's personal sovereignty and freedom to conduct her inquiries as she wishes to burst the bubble and improve her epistemic welfare. Is her interference justified?

On the personal sovereignty model of autonomy, Mary's intervention is clearly unjustified, as Sarah's right of self-determination is more valuable than her epistemic welfare and Mary's interference has infringed this inviolable right. Nonetheless, I believe that this diagnosis of the case is a complete loss. Intuitively, the epistemic gain of being freed from a deeply rooted detrimental condition—such as that of being trapped in an epistemic bubble—outweighs the harm of violating of one's right of self-determination. Moreover, Mary's intervention positively

contributes to improving Sarah's personal autonomy, when conceived along the lines of the model of autonomy as a condition.

For it might well be the case that Sarah displays all the moral virtues that contribute to one's possession of autonomy, but she no doubt fails to display the required intellectual virtues. By omitting their members' exposure to all the available information and counterevidence, epistemic bubbles compromise the members' reliability and good judgment. By favoring interaction among like-minded individuals, epistemic bubbles bootstrap corroboration of the relevant theories or information, thereby deluding their members into believing that they are free to direct their own lives when, in fact, they are only exposed to a limited range of possibilities.

Thus, by bursting the epistemic bubble, Mary not only brings about an epistemic benefit for Sarah, but she also puts Sarah in a position to improve her personal autonomy. Sarah can now (i) acknowledge the functioning of epistemic bubbles and their detrimental effects on one's agency, (ii) inquire into the responsibility of those (if any) people who put her in the bubble and those who actively sustain this structure, and (iii) develop those virtues that are required to become an autonomous person. For these reasons, we can conclude that Mary's paternalistic interference is justified.[81]

Someone might rightly point out that, despite addressing the issue of whether epistemic value can trump one's personal sovereignty, my argument does not solve Bullock's dilemma, in that it remains silent about whether epistemic value can counterbalance considerations of wellbeing. As a matter of fact, nothing in THE CIA BUBBLE elicits the conclusion that this is a case in which Mary's epistemic reasons for intervening clash with the non-epistemic concern with Sarah's overall wellbeing.

As I anticipated early on in the section with the second remark, I concede that becoming more autonomous may negatively affect one's overall wellbeing. For example, it might be the case that, for an individual like Sarah, the non-epistemic value of becoming more autonomous fails to outweigh the psychological costs of getting out of such a longstanding bubble—e.g., the affective costs of questioning the value of important personal relationships, combined with the distress and anxiety that this process generates. Sarah's life is just a mess now that she has realized that the community in which she has been raised was obstructing her access to knowledge and was instilling all this bullshit about the government's corruption and the CIA's bloody activities! No doubt she has grown in autonomy but, in the end, her discoveries did not

[81] It is worth pointing out that a similar argument can be run for interferences with the agency of members of another detrimental epistemic structure, namely *echo chambers* (Nguyen 2018).

change her life for the better, while the psychological effects of such discoveries changed her life for the worse.

This case exemplifies all the relevant features of Bullock's challenge for strict epistemic paternalists endorsing moderate option (a). For the epistemically paternalistic interference brings about an epistemic benefit and improves one's personal autonomy at the cost of violating one's personal sovereignty and conflicting with one's overall wellbeing. Would such an interference be justified?

I would answer this question in the affirmative, for this reason: to the extent that an epistemically paternalistic interference contributes to fostering one's autonomy, it helps the subject interfered with to reach a position at which they can freely choose how to manage their own wellbeing. Thus, although the epistemic reasons in favor of interfering do not align with the non-epistemic considerations of one's wellbeing, they counterbalance this concern by freeing one from any external conditioning on how one should manage their own epistemic and non-epistemic life.

This way out of Bullock's dilemma has two important features: first, it allows strict epistemic paternalists to stick to the idea that epistemic value can trump the concern with one's personal sovereignty while taking into consideration the value of personal autonomy; second, it restricts the range of justified epistemically paternalistic interferences to those interventions that promote the autonomy of the subject interfered with. Thus, on this view, playing physics lectures to a sleeping individual to improve their epistemic welfare will not be justified because the interference, despite allegedly bringing about some epistemic benefit for B, does not contribute to B's overall personal autonomy (see Bullock 2018: 442-443).

6.5 Conclusion

As a conclusion to this chapter, I shall raise a final objection to the proposed solution to Bullock's dilemma. If A's epistemic interference is justified insofar as it improves B's personal autonomy, so the objection goes, then it is not clear whether B had the competence to make decisions about B's own inquiry in the first place. But this undermines the argument I have proposed on behalf of strict epistemic paternalists because it means that the cases considered herein are instances of soft—rather than hard—epistemic paternalism.

We are finally back to the problem I raised in §6.3 about the boundaries of the distinction between soft and hard cases of epistemic paternalism. There are two available conclusions to

our journey. On one conclusion, I have offered an argument that grants justification to a restricted range of cases of hard epistemic paternalism, namely those that, besides improving one's epistemic welfare, foster one's personal autonomy. On the other, I have failed to do so because the cases of epistemic paternalism with which I was concerned are, by Bullock's own standards, cases of soft epistemic paternalism and therefore are already immune to Bullock's dilemma. Either way, the end of the story is a good one: for, besides standard cases in the educational domain, there is a relevant set of scenarios in which our due concerns with promoting one's epistemic welfare can justify our interventions despite the violation of one's personal sovereignty we inevitably incur.

This chapter has concluded the investigation into the rational boundaries of the interaction between non-peer agents and has completed the answer to the Normative Question that I have started to develop in chapters 3 and 5. The next chapter will address the remaining Practical Question.

Chapter 7

Epistemic Authority in Education: The Role of Epistemic Exemplars

Abstract. This chapter explores how epistemically superior agents can help novices acquire intellectual virtues by exercising such virtues themselves or pointing to individuals whom we admire for their agency in the intellectual domain. I call the latter *epistemic exemplars*. The chapter addresses this issue in the context of virtue education and is grounded in Zagzebski's exemplarist theory. After introducing and discussing this theory, I explain how it can serve as a basis for an exemplar-based educational model and discuss its main advantages and most problematic features. Finally, I offer what I take to be the most promising version of an exemplar-based account of education to intellectual virtues and illustrate why and how it departs from the original view.

7.1 Introduction

The investigation conducted so far in the thesis has mostly concerned the transmission of knowledge and understanding from epistemically superior agents to novices. One fundamental epistemic good that has not received due attention yet in the context of the epistemology of expertise is that of intellectual virtues. As I have already mentioned in previous chapters, it can well be the case that an epistemic subject is epistemically superior to another because the former possesses a larger number of intellectual virtues than the latter or one's virtues are more developed than the other's. Authorities of understanding, for example, need to display a wider set of intellectual virtues than authorities of belief, because they are supposed to help their interlocutor acquire understanding in a given domain and this task involves a range of sensitivities and intellectual abilities that those who merely pass along true beliefs to someone who lacks them need not exercise (see §3.6).

What remains to be explored in this context is whether we can expect from epistemically superior agents that they help novices acquire intellectual virtues. In other words, the question that this chapter purports to address is whether a theory of epistemic authority should make room for the idea that helping the novices develop intellectual virtues can be one of the legitimate and desirable services that epistemic authorities (of some sort) fulfill. I shall lay down

my cards on the table from the outset and clarify that I will answer this question in the affirmative. As can be imagined, I will not argue that someone can *pass along* an intellectual virtue they possess to another via testimony or some other form of transmission. Rather, I will argue that someone can *help another acquire* or develop an intellectual virtue by exercising such virtue themselves, if they have it, or by pointing one towards individuals who display and exercise such virtue. Thus, the dynamic that allows a novice to acquire a virtue through the relationship with other people may in some cases involve not only an individual that displays a virtue to some significant degree—which I shall call *an exemplar*—but also an epistemic authority that fulfills an intermediary function, namely that of directing the novice toward the exemplar.

I will say something about this intermediary figure at various points in the chapter, especially as regards the role of the educator or the teacher, but the main focus of my inquiry is *epistemic exemplars*, that is, those individuals who can act as sources of intellectual virtues for people who look at them and their agency in the epistemic domain. The strategy I shall pursue is grounded in Linda Zagzebski's *Exemplarist Moral Theory* (2017) and purports to apply the main exemplarist theses to the domain of intellectual virtues. Thus, this chapter constitutes an overt attempt to build another bridge between the epistemology of expertise and virtue theory as well as to inquire into the implications of this dissertation in the philosophy of education—more specifically, in character education. For, as should be evident, the question about whether an epistemic authority of some sort can help a novice develop intellectual virtues sheds light on a fundamental feature of education, namely its crucial role in the formation of the novice's moral and intellectual character. For these reasons, the chapter aims at addressing the Practical Question that lies at the grounds of this thesis and purports to do so in a way that will hopefully pave the way for further investigations into the relationship between social epistemology and the philosophy of education.

The structure of the chapter is as follows. Section 7.2 summarizes the main features of Zagzebski's exemplarism and explains how it can serve as a basis for an educational model. Section 7.3 offers several remarks about how the model can be extended to intellectual virtue formation as well as an account of epistemic exemplars that aligns with Zagzebski's view of exemplars in the moral domain. Section 7.4 addresses two main objections that can be raised against an exemplar-based account of education to intellectual virtues: one objection worries about the applicability of the model in educational settings, while the other warns us about the risks of indoctrinating the novices while attempting to foster the development of their character traits. Finally, section 7.5 offers what I consider to be the most promising version of an

exemplarist approach to intellectual virtue formation. It is a *liberal* view of an epistemic exemplar, which broadens—in a sense, loosens—the conditions for exemplarity thereby making the whole educational approach more applicable in real-life situations.

7.2 The Exemplarist Dynamic and Virtuous Exemplars

As anticipated, the notion of epistemic exemplars and the exemplar-based account of education to intellectual virtues that I shall offer in this chapter are grounded in Zagzebski's recent exemplarist moral theory. This section illustrates the roots of the view and will therefore mostly concern the moral domain. As will become clear in the next section, this analysis is not a detour from the purposes of the chapter: quite to the contrary, it amounts to a preparatory work essential to understanding the account of epistemic exemplarity that I shall discuss in the remainder of the chapter.

Zagzebski's exemplarist moral theory (2017) is a neo-Aristotelian approach that finds its non-conceptual foundation in the direct reference to morally exemplary figures, identified by the agents through the emotion of admiration. This peculiar foundation marks a neat difference between Zagzebski's exemplarist view and traditional foundational theories, which develop around a concept rather than individuals. On her theory, morally exemplary persons fix the meaning of all fundamental moral terms, such as virtue, good state of affairs, good life, duty, right action, etc. The conceptual machinery upon which the theory develops is Hilary Putnam's and Saul Kripke's semantic theory of direct reference, according to which competent users of a language can pick out instances of natural kind terms such as "gold" or "water" indexically, that is by referring to whatever is the same thing as *that*.

Zagzebski applies this model to the moral domain by arguing that ordinary users of a language can pick out instances of moral terms such as "good persons" by referring to individuals like *that*, that is, moral exemplars we individuate by virtue of their supreme admirability, no matter if we possess or lack well-developed moral concepts. In other words, competent speakers of a language can use certain (moral) terms successfully without necessarily grasping their descriptive meaning. Thus, on this view a *virtue* is what we admire in moral exemplars, a *right act* is what an exemplar would do, and a *good life* is the kind of life an exemplar desires for themselves and others.

Two features of moral exemplarism are particularly relevant for the purposes of this chapter. The first is methodological in spirit: with Zagzebski's own words, the theory is meant

to be "practically useful. By that, I mean both that it gives us directions on what to do and how to live, and it can be used to make us want to do so" (2017: 3). As a result, the theory is particularly interesting from an educational standpoint, in that it should be possible to apply it for helping novices of various sorts (e.g., students and young children) develop virtuous attitudes. However, despite her remarks about the educational advantages of her theory, Zagzebski has not offered detailed suggestions about how we should derive an exemplar-based account of virtue education from her theory.

I shall briefly remedy this lacuna and develop a tool that grants the applicability of this theory to the educational domain, that is, a process of moral development based on the admiration and imitation of moral exemplars according to the principles of exemplarist moral theory. I shall call this tool the *exemplarist dynamic*, and shed light on three main stages: natural admiration, conscientious reflection and emulation.

As regards the first stage, the emotion of admiration should naturally arise before morally exemplary individuals and have the capacity to motivate the subject—that is, to lead them to act in some way—where the resulting action could range from a simple change of attitude towards the exemplar and the object of their deeds to a concrete imitation of their behavior. Needless to say, Zagzebski is aware that things may go wrong in this stage as a result of the fallibility of admiration, as with any other emotion. But if this is the case, why should we trust our emotional dispositions? And how can we do that?

The stage of conscientious reflection helps us address these worries. Zagzebski holds that it is "both inescapable and rational" to naturally self-trust our emotion dispositions (44). It is inescapable because without self-trust in our emotion dispositions, we have no way to figure out whether they tend to produce fitting emotions—that is, emotions appropriately directed towards their intentional object (e.g., fear of a dangerous storm rather than of an innocuous pigeon). It is rational because in the face of the fallibility of our emotion dispositions, the best way to find out whether our admiration towards someone is fitting is to conscientiously reflect on the appropriateness of this emotion—that is, to determine whether our admiration for them survives reflection over time.[82]

When admiration of a moral exemplar naturally motivates us to act and our conscientious reflection reassures us about the rationality of our emotion dispositions, we thereby acquire justification for emulating the exemplar—that is, the third stage in the exemplarist dynamic. In Zagzebski's vocabulary, emulating an exemplar amounts to imitating not only their deeds, but

[82] More on the rationality of trusting our admiration upon reflection can be found in Zagzebski (2012: §4; 2017: §2.5).

also their motive for acting as they do. Thus, emulating the exemplar is a promising way to acquire or strengthen a virtue, for it allows us to form a habitus—that is, to respond to morally relevant situations as an exemplar would do. Emulation does not require that one necessarily replicate an exemplar's actions, for there might well be cases in which their deeds are well beyond our reach. Rather, emulation involves being able to adopt an exemplar's attitude in the situations one finds themselves in, no matter whether they involve decisions with high impact on many people or actions with more limited consequences.

This brief reconstruction of the exemplarist dynamic makes it possible to propose an exemplar-based account of education. The following principles constitute the essential nodes of the account, which have been adapted from Croce and Vaccarezza (2017: 5)[83]:

(i) Human flourishing is the main aim of education, and the acquisition of the virtues is a necessary condition of flourishing (Kristjánsson, 2015: 14);
(ii) Emulation of exemplars is the main way to achieve these aims;
(iii) Emulation as a means of education requires the educator: (a) to elicit the children's admiration via presenting them with genuinely good and imitable models; and (b) to foster the children's capacity for reflection upon prima facie admiration.

Thesis (i) shows that the account is grounded on a combination of moral exemplarism and Aristotelian character education, where priority is assigned to aretaic notions over deontic ones (Steutel and Carr, 1999: 7). It is worth noting that possessing the virtues is not sufficient for flourishing, in that achieving the ultimate aim of moral education also depends on further elements such as "good friends, family, health, basic material provisions, satisfactory education and substantial supplies of 'moral luck' to thrive" (Kristjánsson, 2015: 25). Thesis (ii) illustrates the strategy through which exemplarists attempt to achieve the goal of moral education, thereby accounting for the stage of emulation in the exemplarist dynamic. Thesis (iii) sheds light on the role of the educator within an exemplarist perspective, thereby accounting for the other stages of the dynamic.

The second feature of Zagzebski's exemplarism that is particularly relevant for the scope of this chapter is her account of what (moral) exemplars are. In a nutshell, (moral) exemplars are most admirable individuals because they are paradigmatically good (2017: 20). What we

[83] In the original version, the account focuses on *moral* education but, as will become clear in the remainder of this section, it can apply to intellectual virtues formation as well.

admire in them can be summarized in the following three aspects: (1) the *psychological sources* of their behavior, rather than the external causes that might have contributed to their actions; (2) the *acquired* psychological features that led them to behave as they do; and (3) the *motives of concern for others* they have more than any motives of self-interest (63-64).

On Zagzebski's view, moral exemplars can be distinguished in three categories: *moral heroes* are highly admirable for one specific trait—typically courage—but not necessarily admirable for others; *moral saints* are moved by altruistic concerns to the extent that we expect them to display not only the virtue of charity but all the virtues (96); finally, *sages* are admirable for their extraordinary wisdom (95). As an example of a moral hero, Zagzebski mentions Leopold Socha (2017: §3), whose exceptional courage allowed him to rescue a group of Jewish refugees from the Nazi oppression by hiding them in the sewage canals in Lwów and feeding them for more than a year. As regards sages, Confucius is commonly taken as an ideal candidate. But finding a good example of a moral saint is a very complicated matter given the fact that charity should just be the most striking character trait of a person who is supposed to possess all the virtues.

Thus, although discussing a case of a moral saint in comparison with moral heroes allows us to shed light on the different function that these categories of exemplars can play in education (see Croce and Vaccarezza 2017: 8-10), we should be wary of attributing moral sainthood to human beings and, in particular, with historical figures. This is a recent lesson that we learned with Zagzebski's own example of a moral saint, Jean Vanier, a Swiss philosophy professor who devoted his life to assisting people with severe mental disabilities and who, very recently, has been accused of abusive sexual relationships with six women during his lifetime.[84] For theoretical and educational purposes, then, it might be safer to point to narratives or legends involving figures whose moral sainthood, albeit fictitious, cannot be ruined by historical findings.

Since the analysis of Zagzebski's exemplarism has so far merely focused on the moral character, the reader might rightly wonder why we should think that the considerations offered in this section apply to the intellectual character as well. In the next section, I shall address this legitimate concern and explain how the exemplarist approach can be extended to the domain of the intellectual virtue formation.

[84] See, e.g., https://www.washingtonpost.com/religion/2020/02/23/jean-vanier-once-talked-about-nobel-or-sainthood-candidate-is-accused-abusive-sexual-relationships-with-six-women/.

7.3 Intellectual Virtue Formation and Epistemic Exemplars

That according to Zagzebski the exemplarist theory can apply to intellectual virtues should come as no surprise, if one just bears in mind that, on her view, moral and intellectual virtues share the same nature. In her own words, "an intellectual virtue does not differ from certain moral virtues any more than one moral virtue differs from another … Intellectual virtues are best viewed as forms of moral virtue" (1996: 139). This is a strong claim and, in fact, stronger than we need to justify the extension of her exemplarist approach to the domain of intellectual virtues. All that is necessary to carry out this project is rather that intellectual virtues—no matter whether they can boil down to moral ones—can be acquired through the same strategy. In this respect, it is helpful to point out that, according to Zagzebski, "we admire intellectual virtues in the same way we admire the moral virtues … intellectual virtues are among the acquired excellences … and they belong to the class of traits relevant to exemplarism" (2017: 39-40).

It should be evident from these passages that, on Zagzebski's view, the exemplarist dynamic has prospects to work as a strategy to educate the young to intellectual virtues. The overall framework is completed by these further considerations on epistemic admiration, that Zagzebski has offered in her *Epistemic Authority* book:

> Epistemic admiration plays an important role in acquiring intellectual virtues, learning the norms of inquiry, and assessing our beliefs. We find certain persons and certain behaviors admirable and we learn to think critically and carefully, to be open-minded, intellectually fair, and persevering by imitating persons who have those traits (2012: 90).

This quote sheds some light on the epistemic goods that we can achieve by following the exemplarist dynamic in the intellectual domain. It is nonetheless worth asking ourselves more explicitly: Why should we care about educating the young to intellectual virtues? Or, to put it differently: Why are these virtues so important in our epistemic agency?

The virtue-based theory of epistemic authority that I have developed in previous chapters should have already highlighted that intellectual virtues provide us with the dispositions and motivations to conduct flourishing epistemic activities like passing along knowledge and understanding as well as helping others acquire these goods on their own. To generalize, it could be said that intellectual virtues help us evaluate reasons, acquire knowledge and understanding, help others perform these tasks, and manage our epistemic background, among

other things. In Baehr's words, intellectual virtues help us "become certain types of people—people who wonder, think, ask questions, pursue evidence, listen openly to others, probe for understanding, delight in discovery, embrace intellectual challenges, and stand up for what they believe in" (2015: 35). Not to mention that, in Zagzebski's neo-Aristotelian approach, exercising intellectual virtues is intrinsically good because it makes one admirable and increases one's personal dignity—where this encompasses both an intellectual and a moral dimension.

Now that I have addressed this preliminary question, it is time to have a closer look at the applicability of the exemplar-based model of education to intellectual virtues. The most straightforward strategy to do that is by addressing three central questions about the nature of epistemic exemplarity—a task that I shall complete in this section—and then by investigating some problems that might arise from this model when applied to the formation of intellectual character—a task that I shall complete in the following section.

7.3.1 Natural Talent and Intellectual Virtue

Let us begin with a worry that calls into question the applicability of the exemplarist dynamic in the intellectual domain. One might legitimately suspect that an exemplarist educational approach cannot work in the context of intellectual virtue formation because, as a matter of fact, often we find ourselves admiring intellectual talent or genius that will always be out of reach for most of us.

This worry gets us to an important distinction between natural talent and acquired traits. No educational strategies can allow one to develop a natural talent that one does not possess. In the practical sphere, no trainer will be able to provide me with Usain Bolt's athletic capabilities if they are not already part of my natural setup. Similarly, in the epistemic domain, no teaching methods could let me develop an intellectual talent such as Nikola Tesla's if I lack his genius from the outset. This is because intellectual talents are natural traits that improve one's epistemic activity and run no risk of being lost if one stopped using them for a long time. Those who can solve logical puzzles in the blink of an eye need not practice their ability every week, and those who can spot tiny mathematical errors in extremely complex strings of calculation will keep having such a talent if they are suddenly asked to activate it after years without performing serious mathematical thinking.

In contrast, intellectual virtues are traits that we acquire through exercise, and that need to be cultivated if we care about keeping them (Pritchard 2018: 336-337). For example, if a

member of a research team, call him John, got used to letting his boss disrespect his work and take advantage of it without acknowledging John's scientific contributions, it will become harder and harder for John to exercise intellectual courage and speak up when his boss will perpetrate his unethical conduct again.

A more positive example is offered in C. P. Snow's 1934 novel *The Search*, in which a young and brilliant crystallographer at Cambridge, Arthur Miles, is concluding an experiment that will seal a groundbreaking discovery, one that would boost his scientific career and change the discipline for good. Yet, while analyzing the results of his final experiment, Arthur realizes that he was wrong, in that his x-ray films reveal a patent piece of counterevidence. At that time, anger and frustration make him contemplate the possibility to destroy the counterevidence and publish the results of his work, which would have granted him enough time to secure his own career before the problems with his discovery arose. The novel describes in detail Arthur's struggles and deliberation processes, which resulted in the decision to set aside any pragmatic consideration, stick to evidence and his own love of truth, and therefore announce the failure of his experiments.

As Baehr points out in discussing this story, "Miles is no epistemic saint. He clearly feels the temptation to ignore, distort, even to destroy, evidence that runs counter to his hypothesis" (2011: 143) but, in the end, his actions manifest his intellectual virtues—in particular, his love of truth and intellectual honesty. It is important to note here that it is not just the fact that Arthur decided not to ignore counterevidence that makes him epistemically virtuous. That should be expected of anyone involved in scientific practice—though it is far from obvious when personal, professional, and financial benefits are at stake. Rather, his intellectual virtuousness is revealed by the fact that he habituated himself to cultivating epistemically good conduct in his daily activities as a student first and a scientist later, and this allowed him to resist a strong temptation in the crucial moment.

These examples should suffice to illustrate the difference between natural talents and acquired traits in the intellectual domain. As a conclusion to the discussion of the differences between talent and virtue, two remarks are in order. First, it should be noted that talents and virtues do not go hand in hand: nothing prevents a genius from being intellectually dishonest and, sadly enough, possessing intellectual virtues does not provide one with natural talents. Second, it is the very fact that intellectual virtues are acquired that makes them suitable for the exemplarist approach. For the admiration we may have toward natural talent will not lead us to develop it, but nurturing our admiration toward intellectually virtuous behavior can activate the

exemplarist dynamic and therefore allow us to acquire a new intellectual virtue or develop the ones we already possess.

7.3.2 Who Is an Epistemic Exemplar?

To assess the theoretical plausibility as well as the practical applicability of the exemplarist dynamic in the intellectual domain, it is necessary to have a clear view of who epistemic exemplars are and what features define them.

Zagzebski's own characterization of the kinds of exemplars does not include an explicit account of how these categories can be applied to the domain of epistemic exemplarity—setting aside sages, who are epistemic exemplars given that they excel in an intellectual virtue such as wisdom. Nonetheless, Zagzebski does not close the door to the idea that her list of exemplars can be broadened and further refined[85] and, as I argued in Croce and Vaccarezza (2017: 3-4), Blum's account of heroism and sainthood (1988) appears to fit well with the gist of her view and elicit a broader picture of exemplars. On this view, the main difference between heroes and saints is not just the specific virtue they display, but rather the fact that the former are supposed to excel in one particular virtue, while the latter should display all the virtues to an exceptional level.[86]

The first and foremost advantage of this perspective is that heroism is not anymore limited to the exercise of moral or intellectual courage, but rather to the exercise of one virtue to an exceptional level. More specifically, on this view, it is still required for one to be a hero that one be willing to face risk and danger—thereby making the exercise of courage a necessary component of heroism—but the virtue that one would be admired for need not be just courage. Let me clarify this with the following two examples.

First consider the famous case of Rosa Parks, who left an unforgettable mark in the history of the civil rights movement by refusing to give up a bus seat to a white man who was requesting her to move, in Alabama in 1955. No one will deny that she displayed a great deal of courage in doing what she did, but Josè Medina takes her to be a paradigmatic case of an *epistemic hero* because she shed light on relations of oppression in a community, such as "the invisibilization of certain phenomena, experiences, problems, and even entire subjectivities" (2013: 192). In

[85] Linda Zagzebski has conveyed this in personal communications.
[86] For an alternative account of sainthood, namely one on which a saint is someone who performs supererogatory acts, see Bernstein (1986).

other words, her heroism resides in the combination of her courage to resist oppression and what Medina calls her *meta-lucidity*, that is, an awareness of the effects of oppression in our social structures and the limitations of dominant ways of seeing within our communities (2013: §5). The aforementioned account of heroism provides the appropriate framework to accommodate Medina's account of epistemic heroes, that is, extraordinarily courageous and meta-lucid "subjects who under conditions of epistemic oppression are able to develop epistemic virtues with a tremendous transformative potential" (186).

Those who still think that courage plays too big of a role in Rosa Parks' case to support the proposed view of epistemic heroes might be satisfied with the following example, which is taken from my personal experience with a philosopher whom I had the pleasure to interact with. I shall call him Dr. X. Besides being an outstanding scholar, Dr. X is a fantastic exemplar of intellectual honesty. Whenever he is attending a reading seminar or a philosophical event of any sort and individuates a problem with the argument of the person delivering a talk, Dr. X raises his concerns overtly, without making any distinction of philosophical pedigree or epistemic and practical authority. When it is Dr. X who presents his philosophical work, he is ready and happy to acknowledge the limitations of his own arguments and attribute epistemic credits to those who offer their help. It is with the very same disposition that Dr. X engages with written work: in particular, he pays great attention to the clarity of the sentences, which he explicitly motivates by stating that it is key to the success of the philosophical enterprise at large that we interpret others' views as faithfully and charitably as possible and avoid any risk of ambiguity in our own work.

I had the opportunity to see Dr. X exercise his intellectual honesty on multiple occasions and my initial astonishment became epistemic admiration when I talked with him about the importance of the small gestures through which he exercises this virtue. Now, in some circumstances he could not exercise his intellectual honesty without some courage—think, for instance, of a case in which his interlocutor is a powerful philosopher who does not take well objections and tends to be aggressive in their replies—but, in most occasions, this need not be the case. At least, it need not be the case that his exemplary intellectual honesty brings with it an exceptional degree of intellectual courage.

A further way in which this example supports the proposed account of an epistemic hero has to do with the fact that I also had the chance to notice some respects in which Dr. X is not as virtuous. For instance, his concern with intellectual honesty is so pervasive that he sometimes cannot avoid having arguments with colleagues and scholars whose concern with this virtue does not meet his own standards, thereby proving to lack both practical wisdom and open-

mindedness. On my favored view of epistemic heroes, one's exemplarity regarding a specific virtue need not be accompanied by the presence of other virtues nor by a lack of epistemic vices. On an exemplarist approach, it is an important feature of epistemic heroes that they can be easily associated with a specific virtue they display because educators can quickly direct the novices' attention to that virtue by recalling the hero who possesses it. Furthermore, the very fact that a hero might also display some character flaws—e.g., regarding practical wisdom and open-mindedness in the case of Dr. X—provides educators with the possibility to make the novices work on character shortcomings or epistemic vices, by having them reflect on either the negative effects of the hero's vicious habits or how the hero's struggles to overcome their weaknesses allowed them to become exemplars.

I have not talked much about epistemic saints so far. This is due to at least two sorts of considerations: one has to do with the plausibility of this notion and the other with the theoretical burden that the notion carries with it. As regards the former aspect, I think we can surely come up with narratives featuring the deeds of what we would define epistemic saints, but my reservations about the existence of people who possess all or most virtues to an exceptional degree encompass not only the moral domain but also the intellectual one. While working on the relationship between intellectual virtues, the virtues of the scientists, and the theoretical virtues of scientific theories, van Dongen (2017) characterizes Albert Einstein as someone very close to moral and epistemic sainthood, in that his moral and epistemic virtues qua human being were reflected in his idea of the science and the scientific practice: with van Dongen's words, Einstein's "theoretical and personal virtues here mirrored one another" (2017: 74). I shall stick to my commitment not to enter any analysis of one's sainthood and limit myself to note that, were it to be the case that Einstein's example met the requirements of moral and epistemic sainthood at once, it could constitute the paradigmatic case study for a theory that aims at unifying moral and intellectual flourishing such as Zagzebski's.

The second reason why I have not dedicated much attention to the notion of an epistemic saint has to do with the connection between exemplarity and the unity of the virtues thesis. In the discussion about moral virtues, the notion of a moral saint is put forth by strong unitarists, according to which if an exemplar does not possess all the virtues, she cannot have any of them, (Blum 1988: 51). By contrast, something closer to disunitarism is needed if we care about granting the possibility that one be a moral hero. Luckily enough, nobody in the discussion on intellectual virtues endorses the unity of the virtues thesis as strongly as it is endorsed in the debate about the moral virtues, or so I shall contend. Roberts and Wood claim that there is a unity of intellectual virtues without specifying the details of the view (2007: 80, 310) and Baehr

defends what he calls a *cluster approach* or a *moderate unity*, according to which "character virtues typically are reliable only when possessed in conjunction with other character virtues … Intellectual virtues may necessarily be possessed in certain clusters or subsets" (2011: 65).

The lack of a broad discussion about the unity or the disunity of the intellectual virtues might simply be due to contingent reasons—such as the fact that the virtue epistemology debate is much younger than the virtue ethics one—but another available (and not that implausible) option is that we tend to be fairly liberal about the relationship among intellectual traits, definitely more liberal than we are about the relationship among moral traits. This contention would require more argument than I can provide here, but although some intellectual virtues are closely related to one another, at a first glance it does not seem obvious that for one to be intellectually curious one ought to be intellectually honest too—or for one to be intellectually humble one ought to be intellectually thorough as well.

If this is the case, then it looks as though defending the notion of an epistemic saint becomes less important as a virtue epistemology project, in that the notion of an epistemic hero will allow us to fulfill most of our needs. As I have mentioned early on in this section and extensively argued in Croce and Vaccarezza (2017), from an educational standpoint there are strong reasons to side with disunitarism and concede that one can possess an intellectual virtue to an exceptional degree but lack other virtues. In a nutshell, the main educational advantage of admiring an epistemic saint—namely the fact that we incur no risk of admiring the wrong traits since the exemplar we are looking at has no epistemic flaws—is trumped by the aforementioned advantages of admiring epistemic heroes—i.e., the fact that an educator can immediately refer to a specific virtue by pointing to a hero that excels in that trait and not in others, but also the possibility to have novices work on the flaws and possibly the vices of the hero.

A final advantage of heroes over saints has to do with their imitability. The considerations offered so far should make it clear that heroes are models worthy of imitation for their virtues, but also closer to our imperfection condition given that they also possess non-virtuous traits. In other words, they display what Kristjánsson calls *mixed traits* (2015: 14), whereas saints might look too far away from our experience of imperfect human beings and therefore appear to be less imitable. As a consequence, referring to epistemic heroes rather than saints can avoid—or, at least, limit—the risk that the novice's emotional reaction before an exemplar is one of discouragement rather than a desire to imitate the hero's behavior.

7.3.3 Where Do We Find Epistemic Exemplars?

The final question concerning the notion of epistemic exemplars is rather practical. Supposing that we are happy with the account of epistemic heroes and saints that I have introduced in the previous section, it might still not be clear where we can find these exemplars.

Zagzebski identifies two main sources of exemplars: real life and narratives.[87] Although her discussion focused on moral exemplarity, there seems to be no reason why the same considerations should not apply to epistemic exemplarity as well. Narratives have at least one clear advantage over real life experiences, in that—as Snow's story of Arthur Miles shows—the narrator can describe the process that makes the exemplar become who they are, thereby providing the reader with a detailed description of the path that the exemplar has followed to become virtuous. However, the very same fact that they are a tool at the narrator's—or the writer's—disposal sheds light on the main disadvantage of narratives, namely that they can convey the narrator's own values, interests, biases, and create a distorted exemplary story.

Needless to say, this risk is avoided when we face epistemic exemplars in our real-life experiences. One problem with this source of exemplars, though, is that the opportunity to encounter epistemic exemplars in one's everyday life might be small or entirely dependent on luck. This worry is at least partly due to the fact that Zagzebski acknowledges that the process of becoming virtuous admits of degrees but still defines exemplars as supremely admirable individuals because exceptionally virtuous. Thus, it might well be rare to be in the presence of someone who meets Zagzebski's standards of (epistemic) exemplarity. I will return to this issue in section §7.5; for now suffice it to say that, for an exemplar-based educational approach to be practically useful, we should be willing to broaden the set of exemplars and include at least what I would call *ordinary* or *close-by* exemplars.

The kind of exemplarity that I have in mind here is the one that we might be able to find more easily in our everyday experience and has three main features: *Proximity* in space and time, which brings with it the possibility for the novice to share with the exemplar similar experiences, struggles, and opportunities. *Reachability*, that is the idea that the novice can in principle meet the exemplar and be a first-hand witness of the exemplar's behavior. *Imperfection*, namely the idea that the novice might expect the epistemic exemplar to be exceptional in one (or more) intellectual respect(s) but not flawless, thereby feeling close enough to the exemplar not to be discouraged by their virtuous behavior.

The example of Dr. X that I have introduced in the previous section constitutes a paradigmatic case of close-by epistemic exemplar. He was proximate to my experience, in that

[87] Zagzebski's list includes laboratory experiments that are meant to reproduce the experience we have before exemplary figures and investigate the psychological reactions that they generate. I shall leave this source aside because it concerns moral rather than epistemic exemplarity. For further considerations, see Zagzebski (2017: §3).

we acted in the same working environment for quite some time. For the very same reason he was reachable and the fact that I could establish a personal relationship with him proved to be immensely beneficial, as it made me admire his epistemic behavior and, hopefully, activate the exemplarist dynamic. As I mentioned before, Dr. X was no epistemic saint and discovering this helped me realize that I myself could do better epistemically by trying to behave as intellectually honestly as he did and trying not to behave as close-mindedly as he sometimes did.

7.3.4 Benefits of an Exemplarist Approach to Intellectual Virtue Formation

I shall conclude this analysis of the main features of an exemplar-based approach to intellectual virtue formation with two remarks about the benefits of this model.

First, educating through epistemic exemplars would allow the novice to acquire intellectual virtues in a way that suits their emotional and cognitive needs. This is due to the exemplarist dynamic, whose main advantage is that it builds the conditions for developing character traits based on the novice's expected emotional reactions to the encounter with an exceptionally virtuous figure.

Second, the exemplar-based approach grants an important and multi-faceted role to *intermediary authorities* such as educators, teachers, or parents. On the one hand, it requires that the intermediary authority exercises intellectual virtues in the relationship with the novice(s). I think not only of the sensitivity to their epistemic needs—which I have discussed at length in §3.6—but also of intellectual empathy, which allows the authority to figure out what kinds of struggles the novices might be going through in their intellectual development (as well as interpersonal relationships with other students, if we take into consideration formal educational settings); practical wisdom, which allows the authority to find the most suitable strategies to help the novices go a step further in their process of acquiring intellectual virtues; and such virtues as intellectual humility and open-mindedness, which would allow the educator to assess the effectiveness of the strategies they have proposed to the novices and be ready to revise them if needed.

On the other hand, the exemplar-based approach motivates and encourages educators to exercise their intellectual virtues in the relationship with the novice(s), thereby becoming epistemic exemplars for them. This idea captures a crucial dimension of the approach, namely one that benefits from the refinements I have proposed with the idea of close-by exemplars. Looking at the educator as an epistemic exemplar constitutes an opportunity that has a unique

transformative potential for the novice, because their relationship with the educator is neither confined to some readings nor to a snappy encounter with a person who interacts with them sporadically. Quite to the contrary, novices spend a considerable amount of time with their educators on a daily basis and they do so in the most suitable context, since they already take the educators to be the ones they should be learning from.

These considerations should elucidate the connections between what I have argued in the previous sections about epistemic authorities and the epistemic function of educators. It should now be evident that educators are a particular kind of epistemic authority, namely one whose function might extend beyond providing the novice with knowledge and understanding and encompass the service of fostering the development of their intellectual character. In the best scenario, this function boils down to educators being epistemic exemplars for the novices, but I think it is fair to consider them as *intermediary* authorities in many cases, in that their function includes pointing the novices towards other admirable epistemic agents. It is now time to turn to some crucial objections that threaten the overall plausibility of the exemplar-based account of education.

7.4 Against Exemplarism in Virtue Education

The idea that education, especially formal education, should not only provide students with knowledge and understanding but also put them in a position to develop their character traits looks very plausible in principle, but it quickly loses much of its plausibility as soon as we start thinking how to implement it in educational settings. In this section, I shall concentrate my attention on two worries with any exemplarist approach to intellectual virtue formation, which I shall call the *limited-applicability problem* and the *indoctrination problem*.[88] Let me highlight from the outset that, to a certain extent, the following objections call into question the plausibility of character education at large in both its moral and epistemic dimensions. For the sake of this chapter, I shall focus on the problems they generate for education to intellectual virtues, but the proposed arguments apply to the case of moral virtues as well.

[88] Exemplarism in both its epistemic and moral dimensions has been criticized lately from a variety of perspectives. See, among others, Archer (2019), Szutta (2019), Sullivan and Alfano (2019), Vaccarezza and Niccoli (2019), and Watson (2019).

7.4.1 The Limited-Applicability Problem

The limited-applicability problem is a worry that Alessandra Tanesini (2016) has raised against what Steven Porter calls the *standard approach to virtue formation* (2016). Thus, before addressing the problem in the context of a discussion about intellectual virtue formation, it needs to be shown that an exemplar-based account is part of Porter's standard approach.

According to Porter's analysis, the standard approach to virtue formation deploys several strategies to educate the young to virtue, among which I shall list "(1) direct instruction on the nature and the importance of the virtues; (2) exposure to exemplars of the virtue; (3) practice of virtuous behaviours" (2016: 222). Let me briefly explain how these strategies are supposed to work on Porter's view. Direct instruction amounts to providing novices with a basic vocabulary for character terms—namely, what the virtues are, why they are valuable, and what allows one to distinguish one virtue from the others. This strategy constitutes the most controversial way to foster virtue formation because, in a heavy form, it could inculcate novices with values without providing them with the capacity to rationally assess the values proposed and with the open-mindedness required to question them. Exposure to virtuous exemplars is meant to show novices the virtues in action and allows them to take part in the exemplarist dynamic. Finally, practicing virtuous behaviors is a delicate step, in that an educator is required to foster the rise of circumstances that habituate novices to deploying their virtuous dispositions, and thus to facilitate their emotional and active response to the situations without forcing them to act as a virtuous person would do.

These concise remarks about the above strategies should suffice to explain that an exemplar-based account of education to intellectual virtue fits the standard approach to virtue formation. In fact, it is not necessary that these strategies be deployed in a specific order or proportion for an educational model to be part of the standard approach. Jason Baehr, for instance, considers exposure to exemplars as a subcomponent of direct instruction (2015: 308) yet confers on it a relevant role because it constitutes the most straightforward way to acquaint novices with the virtues. In contrast, exemplarists give a prime role to virtuous exemplars, through which the admiration-reflection-imitation dynamic activates, thereby allowing novices to improve their understanding of concepts pertaining to intellectual character and develop virtuous habits. In this account, the other strategies work as *auxiliary strategies*, in that they support the exemplarist process by fulfilling a function which the dynamic is not in itself suitable to perform.

The limited-applicability problem, as we consider it here following Tanesini's considerations, pinpoints a general limitation of the standard strategies, namely that they only apply to those circumstances in which the students lack non-virtuous attitudes (2016: 524). It is worth pointing out that the objection has classroom settings as its main target: although exemplar-based accounts of virtue education should not limit their scope to children, school settings represent a paradigmatic case study and prime target in their analysis.

Tanesini argues that exposure to exemplars works only with children already disposed to be emotionally moved by exemplary stories and to be motivated to emulate an exemplar's behavior (525). In contrast, this strategy would be detrimental for children who "possess a defensive self-esteem and thus are predisposed toward haughtiness and arrogance" (525) and those who "suffer from self-abasing and obsequious tendencies" (526). The former would overestimate the traits they think they have in common with the exemplar and would disregard what separates them from the exemplar, thereby deluding themselves into thinking they are closer to virtue than in fact they are. For the latter, encountering a virtuous exemplar would amount to further evidence that they are too far away from achieving the virtues, thereby strengthening their negative self-assessment. As a promising strategy to deal with such children, Tanesini suggests self-affirmation techniques involving exercises through which they can reflect on the attributes that define them. For these techniques could reduce "the defensiveness of the arrogant and enhance the explicit self-esteem of the self-abased" (526).

Proponents of an exemplar-based educational model might take Tanesini's point and emphasize that their view need not exclude self-affirmation techniques as an auxiliary strategy to be deployed with arrogant or self-abasing children. Yet whenever children lack tendencies to vices, exposing them to moral exemplars is likely to be more fruitful as an educational method since it can create an opportunity for many students to undertake a process of virtue formation at the same time through the admiration-reflection-imitation dynamic. On careful look, exemplarists can do more: they can acknowledge that for their approach to be educationally effective, the exemplars one looks at need to be tailored to one's possibilities of admiration and imitation. The notion of *close-by exemplars* that I have introduced early on in this chapter points us in the right direction and, as a matter of fact, captures the spirit of the following, perhaps underestimated, remark by Zagzebski:

> Moral improvement comes in stages, and if we aim too high at the outset, we may set ourselves up for failure. Direct imitation of the exemplar may come only after a person has reached a certain level of moral development. Before that, we do better

> at imitating persons who are better than we are, but not so much better that we cannot clearly see the path to becoming like the exemplar (2017: 25).

This idea has been recently tested empirically in the discussion of moral virtues by Han and colleagues, who have shown that "stories of attainable and relevant moral exemplars […] more effectively promoted students' service engagement" (2017: 11) compared with those of historic models. *Attainable* exemplars provide the students with the perception that they can emulate the model without excessive effort, while *relevant* exemplars belong to the students' environment (e.g., family, school, sports team, or community), thereby exercising their virtues in situations familiar to the children. Although the considerations of Han and colleagues focus on moral exemplarity, there seems to be no reason why the account of attainable and relevant exemplars should not apply to epistemic exemplars as well.

One might think the proposed kinds of exemplars go against the spirit of an exemplarist approach, which should focus on authentically exceptional figures. However, since the exemplarist dynamic purports to allow us to imitate epistemic exemplars, as principle (ii) suggests (see §7.2), it becomes crucial to ensure that the proposed models elicit not only our admiration, but also the desire to emulate their behavior, as suggested by principle (iii). Attainable and relevant models may well fall within the set of moral exemplars an exemplarist educational account can deploy, for they fare epistemically better than we do, but in such a way that we do not lose the perception that we can imitate them, and the desire to do so.

How can attainable and relevant models help the exemplarist address Tanesini's objection? I shall contend that interacting with these exemplars can have positive consequences on both arrogant and self-abased students. In both cases, the possibility for the exemplarist to propose a convincing strategy requires a tailored approach, in that the teacher might have to choose different exemplars as suitable for different students. This complicates the educational model from a practical standpoint, but it does not undermine its plausibility and feasibility.

As regards arrogant students, the most promising situation would be one in which the student shares some trait with a virtuous exemplar who also displays intellectual humility. In this case, though arrogance leads the child to think of themselves as similar to the exemplar, it would be possible for them to notice that in several ordinary circumstances the exemplar adopts humble behaviors, that is, acts differently than the child would do. The more relevant a humble exemplar is to the child—i.e., the closer they are to a student's daily life—the easier it is that the child experiences some dissonance between what they think they share with an exemplar and the model's actual behavior. In contrast, it is likely that the student would not be moved by

an intellectually humble exemplar encountered through a narrative, in that the different contexts they live in would make it harder for the arrogant child to be struck by the exemplar's humility and hence to develop a desire to emulate them.

As regards self-abasing students, the prospects of the proposed exemplarist's solution are threatened by the fact that self-abasement negatively affects a child's perception of what is attainable. In such a scenario, it is key to present students with *close-by exemplars*—that is, with relevant, but also epistemically *imperfect* models. For the possibility of seeing them at work in situations familiar to the children and acknowledging both their virtuousness and their flaws would help self-abasing students counteract the distorted perception of the distance between themselves and the exemplar, and develop a motivation to reduce such distance by emulating their behavior.

Nothing I have maintained so far is meant to provide knock-down arguments on behalf of the exemplarists. Rather, I have attempted to show how an exemplar-based account of education to intellectual virtues can address the limited-applicability problem and make room for auxiliary strategies to foster virtue education in less-than-ideal circumstances. In the next section, we shall evaluate whether an exemplar-based account can refute the accusation of indoctrinating children.

7.4.2 The Problem of Indoctrination

It is a shared assumption of most approaches to character education that indoctrinatory educational strategies should be avoided because they inhibit children's capacity for reflection and thereby deprive them of the necessary abilities to freely choose their values and evaluate whether their conduct conforms to such values. A straightforward way to illustrate the problem of indoctrination is offered by this passage from Harvey Siegel's *Education's Epistemology*: "Our educational obligation is not to deliver our students to flourishing lives (or happy ones), but rather to prepare them to consider for themselves (among many other things) the virtuousness of the virtues and their place in their own lives" (2018: 103). Exemplar-based accounts of education are almost inevitably vulnerable to this objection in that they aim at fostering virtue development. The simple fact that educators select epistemic (as well as moral) exemplars based on their supposed virtues already makes the novices prone to being attracted by values that others have chosen on their behalf. Again, with Siegel's own words, these values

"will have been imposed on students from without, rather than having been embraced from within, on the basis of the students' own independent judgment" (86).

Thus, on Siegel's view, the problem lies with the aims of education: rather than promoting *intellectual virtues* as the main goal of education, we should foster the novices' *critical thinking skills*, which promote—rather than impair—their autonomy to choose whatever epistemic (and moral) conduct they consider more conducive to their flourishing. For any approach that "fails to foster students' ability to think critically" turns out to be indoctrinatory, in that whatever "beliefs it imparts, actions it encourages, habits or sentiments it fosters, etc., students so educated will be unable to determine for themselves the worthiness or otherwise of those very beliefs, actions, habits, and so on" (85-86).

The goal of this section is to address the problem of indoctrination on behalf of the proponents of an exemplar-based account of education to intellectual virtues by showing that, pace Siegel, this account can foster students' ability to think critically, thereby having the resources to dismiss the charge. The strategy I shall pursue to accomplish this task amounts to letting the evidence in favor of these claims arise out of a comparison between the critical thinking approach to education and the virtue-based model.

Before delving into this comparative analysis, I shall briefly introduce the notion of critical thinking and discuss its main features. As a first preliminary remark, it is worth mentioning that by sticking as much as possible to Siegel's account, I am restricting the analysis to one of many available educational approaches based on the notion of critical thinking. Yet, it seems appropriate to engage with Siegel's view, in that it is certainly the one that engages most with alternative educational approaches inspired by virtue epistemology.

The notion of critical thinking on Siegel's view is grounded in two main components: the *reason-assessment component*, involving skills and abilities to evaluate the epistemic strength of reasons for an action, a belief, or a judgment, and the *critical-spirit component*, involving dispositions, habits, and character traits that allow one to be disposed to reason assessment and be guided by it in action, belief, and judgment (2018: 21). Siegel is clear that neither of these components, taken independently, can ensure that its possessor displays critical thinking. On the one hand, someone might develop their abilities of rational evaluation to an extremely sophisticated degree and yet lack any disposition whatsoever to deploy them because, for example, they are more inclined to align their thinking styles and deliberation to the beliefs of most members of their community. On the other, someone might have all the appropriate dispositions and character traits required to develop critical thinking without having worked at all on the skills necessary to, say, evaluate the epistemic quality of the evidence at their disposal.

These brief considerations about the structure of critical thinking are accompanied, in Siegel's view, by four reasons why we should consider critical thinking as the most fundamental aim of education (2018: 6-7). The first has to do with the moral principle underlying his view in education and has it that promoting critical thinking is the only way to ensure that students are treated with respect as persons. This is particularly important in the case of young children, in that they lack the capacity to think for themselves: thus, the whole educational enterprise is about providing young students with this capacity. Second, critical thinking puts children in the position to enter adult life because it provides them with self-sufficiency and self-direction. Third, developing and exercising critical thinking skills enables them to learn most—if not all—subjects typical of any formal education curriculum. Finally, critical thinking promotes further abilities like careful analysis, good thinking, and reasoned deliberation, which play a crucial role in the democratic life of our communities and thereby allow them to contribute to their flourishing.

I shall not indulge in commenting on each of these reasons, which certainly get at the crux of what education is about. The interesting question I shall address in the remainder of this section is whether these reasons constitute a prerogative of the critical thinking approach. In other words, I shall inquire into the conceptual distance that separates the notion of critical thinking and that of intellectual virtues. For it is essential to the plausibility of Siegel's critique of virtue-based educational approaches that the critical thinking approach does not fall prey to the same problem of indoctrination.

Siegel admits that intellectual virtues feature in the constituents of the critical-spirit component, in that they are dispositions to obtain epistemic goods, including those concerning the development of a critical spirit. More precisely, Siegel cashes out the relationship between critical thinking and the virtues as follows:

> [Critical thinking] is less than the virtues in that its critical spirit component involves dispositions, habits of mind, and character traits, but that these needn't amount to full-blown virtues; and it is more than the virtues in that its reason assessment component involves all the abilities required to assess reasons in epistemically respectable ways, which the intellectual virtues do not (104).

However, both the distinctions between critical thinking and intellectual virtues that Siegel has suggested have been recently challenged.

For one thing, Jason Baehr has recently taken issue with the idea that critical thinking is *more* than the virtues and has offered two reasons why intellectual virtues also require that the reason-assessment component be satisfied. First, possessing intellectual virtues requires that the subject is able to judge when to engage in activities relevant to that virtue. For example, an open-minded person has the competence to assess reasons in favor of and against considering a competing standpoint. Similarly, an intellectually courageous agent is able to evaluate when it is the case to speak up against their bossy employer and when doing so would be reckless and severely detrimental to their own position. Second, having intellectual virtues requires that the subject aptly and successfully exert the intellectual activity characteristic of that virtue. For example, an intellectually careful individual forms correct epistemic judgments, while an intellectually courageous one makes their voice heard in the appropriate circumstances (2019). Whatever are the necessary intellectual virtues that one has to deploy to display critical thinking skills, it seems clear that fostering these character traits puts students in a good position to become critical thinkers.

For another, the idea that critical thinking is *less* than the virtues can be questioned by appealing to Bailin and Battersby's considerations (2016), according to which the critical-spirit component lacks explanatory force because it fails to take into consideration the issue of the underlying motivation for action. Possessing critical dispositions is no guarantee that one has the right motivation towards the domain of inquiry. For one could even develop a critical spirit as a result of an indoctrinatory relationship with a past teacher and this would obviously clash with the very idea of critical thinking. Instead, understanding the critical-spirit component in terms of intellectual virtues avoids this paradoxical consequence, in that the motive underlying the virtues involves an appreciation of the value of virtuous behavior and of how cultivating that behavior could lead to intellectual flourishing. These considerations show that, far from being a flawed epistemic value of education that makes us prone to being indoctrinated, intellectual virtues are in fact an important shield against indoctrination. Thus, pace Siegel, critical thinking should not be less than the virtues as regards the disposition or critical-spirit component: in fact, we would do better to consider it in terms of intellectual virtues.

The discussion about the relationship between critical thinking and the intellectual virtues could go well beyond these considerations, which should nonetheless suffice as a response to Siegel's claim that virtue-based approaches of education to intellectual virtues are bound to indoctrinate the young.[89] As I have just shown, this educational model can be indoctrinatory

[89] Further arguments in favor of a critical thinking approach over a virtue-based one have been proposed in recent work by Kotzee, Carter and Siegel (2019). I plan to engage with this paper in future work.

only insofar as it forgets its own mission and fails to promote the development of the right motivation towards one's intellectual activities.

Furthermore, it should be noted that an exemplar-based account of education constitutes the least problematic version of a virtue-based educational model. For this account makes minimal use of direct instruction as an educational strategy and thereby reduces the risk of being accused of indoctrination. Within this approach, direct instruction merely provides the students with a basic vocabulary about the character traits they could discover in the exemplars. But it is the exemplarist dynamic that does the main work by triggering the students' *natural* emotion of admiration towards the exemplars under consideration.

Someone might worry that the proposed argument works with Siegel's account but fails to apply to other critical thinking approaches to education. As a final remark to the argument of this section, I want to mention another prominent view of critical thinking, offered in Robert Ennis' work, which appears to be even more favorable towards intellectual virtues than Siegel's own view. According to Ennis, critical thinking is "reasonable reflective thinking focused on deciding what to believe or do" (2018: 166). Interestingly enough, his list of *ideal dispositions* of a critical thinker (167) includes such elements as "be open-minded", "seek as much precision as the subject admits", and "use credible sources and mention them". It should be clear that Ennis is basically after such intellectual virtues as open-mindedness, intellectual thoroughness, and intellectual honesty.

Furthermore, among his proposed strategies to teach critical thinking, Ennis suggests that the educator sometimes asks why-questions both when she agrees with the students and when she does not and that she asks them to address a question to which she does not know the answer or whose answer is controversial to discuss it with them. Both strategies pick up again on important intellectual virtues: the former strategy appeals to the virtue of inquisitiveness, which has been explored in recent work by Lani Watson (2015, 2016, 2019), while the latter requires the educator to exercise her intellectual humility and intellectual honesty. As should be evident, these strategies show not only that the intellectual virtues play a central role in Ennis' critical thinking approach to education, but also the importance of the educator *qua epistemic exemplar*, that is, of someone who exemplifies the virtues in a daily—or, at least, an ordinary—relationship with the students and in a context with which they are familiar.

The considerations offered in this section demonstrate that the exemplar-based account to education—qua a specific kind of virtue-based educational approach—can successfully address the problem of indoctrination.

7.5 In Favor of a Liberal Account of Epistemic Exemplarity

What I have done so far in the chapter is to delineate the contours of an exemplar-based account of education to intellectual virtues, motivate the view, and defend it from a couple of serious objections that have been recently raised in the epistemology and education literature. In this section, I shall go beyond the standard exemplarist approach and argue that, for what matters educationally, we have good reasons to endorse a *liberal* account of epistemic exemplarity, namely one on which epistemic exemplars need not always be virtuous or imitable. I have defended a very similar view of moral exemplarity elsewhere (Croce 2020a) but I am convinced that the crux of the view can apply to the intellectual domain as well.

As a preliminary remark, I want to address a concern with the overall spirit of the project: one might worry that departing from Zagzebski's original account of exemplars betrays her vision of the educational implications of the exemplarist theory. Let me offer two reasons why I think we should not worry about this problem.

First, since our project aims at assessing how exemplars can be used in educational settings to foster virtue development, it might turn out that educational effectiveness trumps orthodoxy regarding the exemplarist theory. It is to be expected that theoretical frameworks require adjustments whenever we want to apply them to real-life situations: even more so when a rather ideal theory of moral and epistemic exemplarity has to be implemented in non-ideal settings like ordinary classrooms.

Second, were it to be the case that the liberal account I shall put forth waters down the notion of an exemplar to the point that it is no longer clear whether we can still talk about an exemplarist approach to virtue education, this need not have any consequences at the level of the theory. Proponents of exemplarism could still hold firm to an ideal understanding of exemplarity and the exemplarist dynamic for theoretical reasons, while conceding that in non-ideal situations some improvement can arise out of an encounter with less-than-exemplary figures. In short, the very fact that there might be reasons for switching to a liberal notion of an epistemic exemplar from an educational standpoint is not in itself revealing of any flaw in the exemplarist theory.

The alternative account of an epistemic exemplar that I propose liberalizes Zagzebski's exemplarist theory in two specific respects. On the one hand, I shall challenge the exemplars' virtuousness requirement, that is, the idea that exemplars are paradigmatically good individuals, and show that in some circumstances less-than-virtuous agents can well serve the purposes of

epistemic exemplars. On the other, I will take issue with the third step in the exemplarist dynamic, that is, the idea that emulation is the main way for novices to develop the virtues, and argue that some admirable individuals whom we either cannot or do not wish to imitate should nonetheless be regarded as epistemic exemplars.

7.5.1 Epistemic Exemplarity without Virtue: Enkratic Exemplars

The first non-standard case of epistemic exemplarity I want to propose is represented by what I will call *enkratic exemplars*. The *enkratês* in Aristotle's ethics (NE VII) is someone who acts in accordance with reason but experiences a feeling that is contrary to reason. As such, *enkrateia*—i.e., continence—is worse than virtue in that the latter involves no internal conflict within the agent to act in accordance with reason. It is precisely for this reason—i.e., the fact that the *enkratês* struggles to do the right thing—that enkratic exemplarity could constitute a decisive step in virtue development on an exemplarist approach, or so I shall argue.

It is worth stressing that I am investigating the possibility to extend the standard notion of moral enkrasia to the domain of inquiry or intellectual agency. Thus, my considerations about *epistemic enkrasia* still fall within an Aristotelian framework that is concerned with the study of human flourishing. The more specific notion of epistemic enkrasia that is discussed in epistemology—according to which one should avoid adopting a belief state that one takes to be irrational (e.g., Askell 2018; Lasonen-Aarnio 2020)—constitutes just one of the several dimensions of enkrasia in the domain of intellectual agency.

Let me offer two cases of enkratic exemplars which could suggest how these figures might be deployed in educational settings. On the one hand, an educator might present the novices with a tale like the following: Paul is a young scientist working in a research lab that is developing antibody home test kits to identify whether people have been infected by the Covid-19 virus. His current task is to double-check that the kits which are about to go on the market respect the safety and manufacture requirements. This task is boring, and Paul feels the desire to skip some checks as he knows that all the kits that have been prepared by the lab so far fulfilled the requirements. Despite struggling with himself, he commits to staying focused and doing his checks accurately, as this is crucial to the safety of his community. On the other hand, opportunities for shedding light on enkratic exemplarity might easily arise out of little episodes in the everyday life of a classroom. For example, a student who admits that her classmate is right and she herself is wrong in a discussion about the main causes of dinosaur extinction

despite feeling the impulse not to give up in the discussion and relentlessly stick to her guns would amount to a good case of an enkratic exemplar in an ordinary setting.

As should be evident, *enkrateia* cannot be a mark of epistemic exemplarity for someone who already knows how to master passions and desires when they conflict with reason. Those who are always intellectually perseverant and thorough in their activity and those who never have the impulse to stand out at any rate in a discussion will not see anything special—more precisely, admirable—in Paul's and the student's behavior. However, the epistemic population encompasses other categories of individuals, such as vicious agents, *akratic* individuals—i.e., those who experience a feeling that is contrary to reason and who often act in accordance with that feeling despite being aware that they are going against reason—and children who are approaching epistemic considerations about the relationship between passions and reason for the first time. My contention is that for these categories, enkratic individuals might constitute a legitimate kind of epistemic exemplars, if not the most appropriate one.

What makes them suitable for epistemic exemplarity from an educational standpoint is the particular combination of admirability, imperfection, and the struggle they display. Their imperfection is evident in the fact that acting well is neither natural nor easy for them, but at the same time, it is also a clue of relevance and attainability, to stick with the terminology of Han et al. (2017). Moreover, their struggle provides the novice with an opportunity to reflect on the interaction between reasons and passions, starting with exemplars whose actions are likely to be much closer to the novice's everyday experience than the typical narratives of epistemic heroes and saints.

One might worry that this whole idea of enkratic individuals as epistemic exemplars goes against the exemplarist approach in that the admiration towards them cannot be fitting if we just remind ourselves that they fall short of virtue. Let me try to appease this critic with two considerations. First of all, to fulfill the admiration stage in the exemplarist dynamic we need *motivating* admiration, that is, an emotion that generates the desire to imitate the exemplar and activates positive attitudes and reasoning about their virtuousness. But as I mentioned in §7.3.2, it is far from clear that motivating admiration is what we can expect from novices when we present them with epistemic heroes and saints: quite to the contrary, there is a fair chance that admiration for those figures will backfire and actually discourage the novices.

Secondly, there is surely something admirable about the behavior of enkratic individuals: first, they act well; second, they do so out of good motives; third and most importantly, they do so as a result of an inner struggle. This shall suffice to conclude that whenever a child, an akratic agent, or a vicious one admires an enkratic exemplar, their admiration fits its object. A similar

verdict has implications for Zagzebski's cognitivist view of emotions (2003). For one thing, it suggests that an emotion's fittingness might depend not only on the feeling and the object towards which the feeling is directed, but also on the agent's perspective—in the case of admiration, on their moral and epistemic position. This explains why admiration towards enkratic exemplars is fitting for the aforementioned categories of agents, but not for those who are at later stages in virtue development. For another, it suggests that an emotion's fittingness might admit of degrees, with enkratic exemplars sitting near the low end of the spectrum and epistemic—as well as moral—saints near the high end.

The proposed liberalization of epistemic exemplarity also has wider implications for the exemplarist dynamic. As regards the third stage, it goes without saying that the imitability of enkratic exemplars is somewhat limited, as there is nothing in their behavior that those who already master passions should emulate. Yet, emulating enkratic exemplars could constitute a reasonable path for vicious and akratic individuals to abandon their bad habits and move closer to epistemic virtue. Furthermore, regardless of their imitability, exposure to enkratic exemplars can lead novices, akratics, and vicious agents to increase the awareness of their limitations and start questioning their behavior: no doubt this constitutes an essential building block in the path towards becoming virtuous. This analysis of enkratic individuals made a case for the idea that less-than-virtuous agents can count as epistemic exemplars and, more generally, that the educational function of epistemic exemplars goes beyond the standard model of virtue development by emulation.

7.5.2 Epistemic Exemplarity without Emulation: Injustice Illuminators

The second non-standard case of epistemic exemplarity I shall offer puts pressure on emulation as the third stage of the exemplarist dynamic and is constituted by what I shall call *injustice illuminators*. In short, I conceive of injustice illuminators as the broad genre of which Medina's epistemic heroes constitute the supremely admirable species. Before accounting for why a theory of epistemic exemplarity should make room for injustice illuminators, I shall clear the ground of a preliminary worry and explain why I have decided not to stick to Medina's original label to identify those who excel in meta-lucidity. In particular, I shall offer two reasons in favor of this choice, which is more conceptual than terminological in spirit.

The first reason is that injustice illuminators need not be epistemic *heroes*. As we have seen in §7.3.2, epistemic heroes like Rosa Parks are extraordinarily courageous and meta-lucid

individuals who manage to act according to virtue in dramatic situations of epistemic oppression. However, an injustice illuminator can well be a student or a colleague who manages to cast light on some sort of small-scale discrimination that has been perpetrated in the classroom or workplace, yet it is not required that their virtuous intervention has the "tremendous transformative potential" Medina has in mind when he talks about epistemic heroes (2013: 186). Thus, while epistemic heroes are injustice illuminators by definition, the opposite is not true.

The second reason is that the contribution of injustice illuminators goes well beyond the epistemic domain and therefore considering them as mere epistemic exemplars will not do justice to these individuals. I cannot go further into whether epistemic heroes can count as moral exemplars: my inclination is that they do so, but Medina is clear that the contribution they give pertains to the epistemic domain. All I need to make my point about injustice illuminators is to highlight that their contribution towards calling out instances of epistemic—and, in particular, testimonial—injustice (Fricker 2007) encompasses both an epistemic and an ethical dimension. As Fricker clearly points out (122-ff.), the phenomenon of testimonial (in)justice is hybrid in nature, in that it has to do with (a lack of) truth and (in)justice at the same time. Thus, there seems nothing wrong in considering injustice illuminators as both as epistemic and moral exemplars insofar as their acts contribute to counteracting unjust dynamics or restoring just practices within a given community.

Having clarified why we should not reduce the notion of an injustice illuminator to that of epistemic heroes, we can now explain why a theory of epistemic exemplarity should make room for injustice illuminators. As I have done for enkratic exemplars, I shall argue that the behavior of injustice illuminators is both admirable in some respects and conducive to the intellectual development of the novices who admire them.

The activity of injustice illuminators can be unpacked as follows: first, they have the ability to see what is epistemically unjust about a given situation, relationship, or social structure; second, they realize how members of a community might ignore the problem—that is, by failing to notice the wrong—if not be part of the problem themselves—e.g., by being those who perpetrate the wrong; and third, they find a way to reveal such wrong, to make it evident in the relevant community. Doing so requires both courage and meta-lucidity, that is, two character traits for which injustice illuminators are worthy of our admiration. However, there is a clear limit to their imitability, in that one cannot simply choose to be part of an oppressed group and therefore one cannot decide to emulate the deeds of those who are sensitive to injustices and discrimination because they have suffered (or are suffering) them.

Thus, one might worry that injustice illuminators cannot be epistemic exemplars if the emotion of admiration we might experience towards them fails to elicit emulation—at least for those members of a community who do not belong to an oppressed minority. In fact, this should be the verdict issued by a traditional exemplarist approach. A liberal account, though, can accommodate the idea that injustice illuminators are epistemic exemplars insofar as the encounter with these figures motivates those who admire them to reflect on the problem highlighted, changing their attitudes towards the victims of injustice, and elaborating ways of counteracting the injustice. Although, strictly speaking, only other members of oppressed groups may emulate them, what makes them exemplars is the contribution they make towards fostering human flourishing. They might not allow us to develop the virtues they display as we would expect from models of what I shall call *direct exemplarity*, but they still allow us to get closer to moral virtue, as models of *indirect exemplarity* can well do.

An opponent of the liberal account of epistemic exemplarity might want to object that the proposed distinction between direct and indirect forms of exemplarity just is an *ad hoc* maneuver to avoid the evident problem that injustice illuminators do not fit the exemplarist dynamic. As a reply to this objection, I shall point out that support for the idea of indirect exemplarity can be found in Archer's (2019) recent account of admiration. As he clearly argues, Zagzebski's emulation view of admiration—i.e., the thesis that admiration is meant to provide motivation for emulating the object of admiration—is problematic: on the one hand, it is committed to the idea that the object of admiration can only be human beings, while there seems to be no principled reason why we could not also admire nonhuman animals and objects;[90] on the other, it unduly restricts the range of appropriate responses to admiration to a desire to emulate the object of admiration. As a more promising alternative, Archer puts forth and defends the value promotion view of admiration, according to which "in prototypical cases of admiration, an agent who experiences admiration will be motivated to promote the value(s) that they admire in the object of their admiration" (146).

Once we acknowledge that "a desire to emulate is one of several ways in which admiration typically motivates" (148), it becomes evident that indirect exemplars and, in particular, injustice illuminators can be legitimately considered as epistemic exemplars. As I have argued, admiration for them promotes a range of epistemically—as well as morally—valuable achievements such as raising one's sensitivity to epistemic (and moral) injustice and eliciting counteracting attitudes to prevent the perpetration of unjust social dynamics. Furthermore, as Sherman (2016) has it, the process of changing attitudes towards the victims of injustice and

[90] See also Korsgaard (2019) on the distinction between human beings as exemplars and exemplary things.

elaborating ways to counteract it requires developing and deploying other virtues, such as vigilance, humility, and conscientiousness. These considerations motivate the conclusion that injustice illuminators can count as epistemic exemplars and therefore provide support for a liberal account of epistemic exemplarity, as injustice illuminators amount to a case of non-imitable exemplars.

7.5.3 Benefits of a Liberal Account of Epistemic Exemplarity

Let me summarize the main take-home messages of the liberalizing project I have undertaken in this section and stress the implications of the proposed arguments for the exemplarist dynamic. The goal of my project here was to signal two directions in which the standard account of epistemic exemplarity should be liberalized if we care about its applicability in educational contexts. On the one hand, I showed that intellectual virtue is not a necessary requirement of epistemic exemplarity, as enkratic individuals can be exemplars for young children as well as akratic and vicious adults. On the other, I showed that imitability is not a necessary feature of epistemic exemplars, in that injustice illuminators can be considered exemplars despite the fact that most of us will never be in a position to emulate their behavior.

My proposal calls for a liberalization of the exemplarist dynamic, for considering emulation as the only appropriate motivational response to admiration is unduly restrictive. The liberal approach is more promising in that it allows us to be pluralist about the ways in which the encounter with epistemic exemplars can promote epistemic value. Whenever emulation is not a realistic opportunity, exposure to epistemic exemplars may still allow novices to change their attitudes towards some epistemically relevant circumstances or find ways of counteracting injustice. Both activities are in fact likely to promote virtue development as not only Zagzebski but also proponents of character education expect from an exemplarist approach to intellectual virtue formation.

As a final remark, let me stress that the liberal account I have defended herein has an important virtue, namely the fact that it is not just liberal from a theoretical standpoint, but it is also *liberating* from an educational perspective. For starters, the idea that emulation should not always be the expected outcome of a novice's interaction with epistemic exemplars frees novices from the ambiguity of not knowing whether they should become able to replicate the exemplar's deeds. The account further reduces the pressure on novices by showing that a realistic path towards human flourishing might require intermediate steps in which they should

be exposed to enkrasia rather than exceptional virtue. Finally, the account is also liberating for educators, who need not necessarily present novices with almost-flawless exemplars and then deal with the side effects of an admiration that backfires, but who can rather resort to more manageable exemplars insofar as they are relevant, attainable, and sometimes even less-than-virtuous characters.

As anticipated at the beginning of the section, nothing I have argued here commits Zagzebski and proponents of standard exemplarism to give up on any of their theoretical cornerstones. That said, I am also convinced that my arguments suffice to motivate the need for a more plausible and effective educational application of the exemplarist theory.

7.6 Conclusion

This chapter has offered an inquiry into how epistemically superior individuals can help novices acquire intellectual virtues. By addressing this question, I have attempted to fill a gap in the current debate in the epistemology of expertise, which to date has only investigated the transmission of knowledge and understanding between epistemically unequal individuals. The strategy I pursued amounted to exploring the notion of epistemic exemplarity and its role in a paradigmatic context for any theory of epistemic authority, namely education. This project required an in-depth engagement with Zagzebski's exemplarist theory (§7.2) as well as an attempt to outline an educational approach to foster intellectual virtue formation that respected her theoretical requirements (§7.3) without falling prey to relevant objections (§7.4). As section 7.5 revealed, I had further reservations about the applicability of an exemplarist approach in concrete educational settings and I have addressed them by offering a liberal version of the account, thereby indicating what I take to be the most—but also the only—promising direction for exemplarism in virtue education.

The overall project of investigating the plausibility of an exemplar-based approach of education to character virtues is far from being concluded. For one thing, it ought to be assessed whether a liberal account of virtue education withstands—or, at least, accommodates—the challenges moved by opponents of a neo-Aristotelian version of exemplarism (e.g., Sullivan and Alfano, 2019). For another, the approach itself—at least, in the version discussed here and inspired by Zagzebski's work—is relatively new: thus, it is to be expected that the discussion will grow in the next few years and continue to attract the interest of ethicists, epistemologists, and education theorists.

Nonetheless, I believe that this chapter has offered a twofold contribution. On the one hand, it has outlined a novel view in the philosophy of education in sufficient detail, thereby providing those who are interested in the topic of epistemic exemplarity with a basis for comparison. On the other, it has addressed the Practical Question that lies at the grounds of this thesis and has shown that, to some extent, the relationship between epistemic authorities and novices can constitute an opportunity for the former to exercise their intellectual virtues and act as an epistemic exemplar for the latter or, at least, help novices acquire intellectual virtues by pointing them towards exemplary people.

Chapter 8

Conclusion

This thesis has undertaken a journey in the lands of social epistemology to explore how a relationship among epistemically unequal agents can be beneficial to the epistemically inferior one. The project was initially motivated by a longstanding interest in the theme of epistemic authority and expertise, namely one that had primarily to do with a recurring first-person question: Whom should I trust? Many—often daunting—events and societal developments have quickly made me realize the relevance and the urgency of an inquiry into the epistemology of expertise for our communities at large. The timing of this work turned out to be fortuitous, though the pace at which the debate on this theme is growing might cause some of the considerations offered in the thesis to age quickly. Nonetheless, this is for the better. For it will be a clear sign that philosophers have not wasted the opportunity to do what they can to counteract the widespread climate of distrust in expertise that we currently experience in so many domains of our everyday life.

Interestingly enough, the flow of the most recent events has—at least temporarily—prompted us to stop talking about the so-called "post-truth" atmosphere (Fuller 2018; Haack 2019) that was probably receiving more attention than it actually warranted. We still wake up every morning with new conspiracy theories and pieces of fake news, but it looks as though the outbreak of Covid-19 has made it clear again that our lives depend on what we know and understand about the world we inhabit and therefore on experts who devote their lives to the epistemic progress of our communities. Perhaps, the word of scientists is regaining some of the epistemic authority that it had lost in the recent past—or, at least, this is my (optimistic) impression regarding the current European situation. If this were true, it would undoubtedly be a positive fact, though the price our communities are paying for this is of course too high. I do not know whether the reputation of experts will fluctuate in the coming months or years along with the development of this and future pandemics, but everyone must do their part to ensure that we do not lose the level of social awareness of the importance of experts that we have just started to appreciate again.[91]

[91] I tackle some of these issues about experts and the Covid-19 pandemic in Croce (2020c).

This might sound overly optimistic or idealistic, but I like to think of my thesis as a contribution in this direction. More specifically, I have attempted to address three main questions about epistemic authority which were much in need of further investigation.

The Conceptual Question prompted me to clarify what it takes for someone to be an expert in the cognitive domain. I have shed light on two main conceptual aspects in the epistemology of expertise. On the one hand, in chapter 2 I have disentangled a longstanding ambiguity in Alvin Goldman's seminal account of an expert by offering reasons to distinguish the authority of those who contribute to the epistemic progress of a discipline—which I have proposed to call *experts*—and those who help novices improve their epistemic standings in a given domain—which I have called *epistemic authorities*. On the other hand, in chapter 3 I have introduced a framework for individuating various kinds of experts and authorities based on the services they can perform and the intellectual virtues they are required to deploy. This virtue-based framework is particularly helpful because it can make room for basically any other form of epistemic authority that we might encounter, and it allows us to figure out what epistemic features someone ought to possess to fulfill a particular function in the epistemic goods market. Chapter 4 has offered an initial application of my theory of epistemic authority to the domain of collective epistemology.

The Normative Question allowed me to inquire into the rational commitments and requirements of testimonial exchanges between epistemically unequal agents. In chapter 3, I have defended a *limited-preemptionist* view of epistemic authority which tells us that we should set aside our opinion and defer to expert testimony whenever we disagree with someone whom we have good reason to consider epistemically superior to us but also when we have reason to think that our interlocutor has already taken into consideration all the evidence at our disposal. Chapters 5 and 6 offered an in-depth analysis of epistemic paternalism, investigating the extent to which epistemically superior agents can interfere with the agency of the novices for the epistemic welfare of the latter without having their consent. In chapter 5, I have offered an account of what it takes to be a paternalist interferer on my virtue-based theory of epistemic authority. In chapter 6, I have discussed the alleged incompatibility between the value of personal autonomy and the epistemic value brought about by the paternalistic interferences via addressing Bullock's dilemma against the legitimacy of epistemic paternalism.

The Practical Question provided the opportunity to discuss the educational implications of the current debate in the epistemology of expertise. My main goal was to shed light on an underestimated aspect of the relationship between epistemically unequal individuals, namely the possibility that the novice acquires or improves their intellectual virtues via their

relationship with an epistemically superior interlocutor. Education naturally constituted the most suitable context in which to discuss this issue, because it is certainly key to several—if not most—educational approaches that the novices develop their epistemic abilities and intellectual character traits. To navigate this issue, I have appealed to Zagzebski's recent exemplarist theory, the educational implications of which I had explored elsewhere in the context of moral education. Chapter 7 offered an extended analysis of what it takes for one to be an epistemic exemplar, that is, someone whom a novice can admire and imitate, thereby developing virtuous traits. The novel exemplar-based account of education to intellectual virtues I have outlined and defended builds a bridge between the epistemology of education and standard neo-Aristotelian character education.

Although this thesis could only offer a sample of the applicability of my theory of epistemic authority to fundamental questions in the epistemology of expertise, I am convinced that its potential for application is wide enough to cover further "hot" topics within the current debate in social and applied epistemology. As a final remark on the thesis, I want to briefly point out how this work on epistemic authority can be relevant for some of these issues—on which I plan to work in future research.

First, below the surface of the epistemology of expertise which I have attempted to scratch in this thesis, there are at least two further questions that a comprehensive theory of epistemic authority has to address. On the one hand, it seems plausible to imagine that epistemic inequality could also bring about epistemic—and other—benefits for the epistemically superior agent. Besides the fact that sharing one's knowledge or understanding with novices can help an expert refine their own epistemic background, it should be expected that reliably fulfilling one's epistemic function helps one strengthen and develop one's abilities and intellectual virtues. The issue of what experts and authorities can gain from devoting their time to interacting with novices is worth further inquiry because, among other things, it would motivate scientists and academics to improve their outreach skills. On the other hand, it looks as though a theory that seeks to make room for various kinds of epistemic authorities should also account for the role of those agents who disseminate specialist knowledge among the wider public, thereby making accessible information that could otherwise be unhelpful for most members of the epistemic community. My initial hunch is that disseminators are neither experts nor mere epistemic authorities: as in the case of virtuous paternalist interferers, there is something specific to the function of epistemic disseminators, or so I plan to argue in future research.

A second direction in which my theory of epistemic authority can be broadened has to do with the nature of the relationship that connects the two parties in a testimonial exchange among

unequal agents. The Normative Question has told us how we *should* rationally behave before expert testimony, but it seems no less important to find out how we *can* so behave. The epistemology of expertise would largely benefit from further investigation into the philosophy—better, the epistemology—of trust and, in particular, into how a novice can be a good or a virtuous truster (Carter 2019). In the context of an inquiry into the ongoing relationship between unequal epistemic parties, it would be helpful to build another bridge and connect the discussion on epistemic authority with the debate on the epistemology of friendship and epistemic partiality (e.g., Goldberg 2019; Hawley 2014; Kawall 2013; Stroud 2006). For though novices need not befriend their epistemic authorities, it is not hard to imagine that some stable relationships between epistemically unequal individuals might generate normative questions similar to the ones that arise in the discussion of the epistemic and moral demands of friendship.

A third and final direction in which this work on epistemic authority could be deployed has to do with the notion and the function of expertise in other domains, such as morality and politics. Despite many overlapping considerations, the metaethical debate about moral experts and the (meta)epistemic debate about epistemic authority tend to be run in parallel. A virtue-based framework appears to provide an interesting opportunity to bring together these discussions and explore the implications that each topic may have on the other. To my knowledge, this is a project that is still very much in need of further exploration.

As these brief considerations suggest, the research I have conducted in this thesis in the epistemology of expertise can have a wide impact within other areas of philosophical interest and beyond, particularly as regards the political and moral conduct of our communities. I hope this work has contributed to showing what experts and epistemic authorities look like and why we need to trust them—today more than ever. But I also hope that it has opened enough windows and built enough bridges that other scholars can make use of it to further advance the discussion on epistemic authority and expertise. The epistemic challenges we currently face are of such serious proportions that only a collective effort can have any prospect of succeeding.

Bibliography

Ahlstrom-Vij, K. (2013). *Epistemic Paternalism: A Defence*. Basingstoke: Palgrave Macmillan.

Annas, J. (2011). *Intelligent Virtue.* New York: Oxford University Press.

Antony, L. (2018). "Quine as Feminist: The Radical Import of Naturalized Epistemology." In L. Antony and C. Witt (eds.), *A Mind of One's Own*, London: Routledge, pp. 110-153.

Antony, L. (1995). "Sisters, Please, I'd Rather Do It Myself: A Defense of Individualism in Feminist Epistemology." *Philosophical Topics* 23(2): 59-94.

Archer, A. (2019). "Admiration and Motivation." *Emotion Review* 11(2): 140-150.

Askell, A. (2018). "Epistemic Consequentialism and Epistemic Akrasia." In K. Ahlstrom-Vij and J. Dunn (eds.), *Epistemic Consequentialism*, Oxford: Oxford University Press, pp. 290-310.

Baehr, J. (2011). *The Inquiring Mind: On Intellectual Virtues and Virtue Epistemology.* New York: Oxford University Press.

Baehr, J. (2015). *Cultivating Good Minds. A Philosophical and Practical Guide to Educating for Intellectual Virtues.* Long Beach.

Baehr, J. (2019). "Intellectual Virtues, Critical Thinking, and the Aims of Education." In M. Fricker, P. Graham, D. Henderson, N. Pedersen, and J. Wyatt (eds.), *The Routledge Handbook of Social Epistemology*. London: Routledge, pp. 104-121.

Bailin, S. and Battersby, M. (2016). "Fostering the Virtues of Inquiry." *Topoi* 35: 367-374.

Bernstein, M. (1986). "Moral and Epistemic Saints." *Metaphilosophy* 17(2-3): 102-108.

Bird, A. (2010). "Social Knowing: The Social Sense of 'Scientific Knowledge.'' *Philosophical Perspectives* 24(1): 23-56.

Blum, L. (1988). "Moral Exemplars: Reflections on Schindler, the Trocmes, and Others." *Midwest Studies in Philosophy* 13(1): 196-221.

Boyd, K. (2017). "Testifying Understanding." *Episteme* 14(1): 103-127.

Boyd, K. (2019). "Group Understanding." *Synthese*. doi: 10.1007/s11229-019-02492-3.

Brady, M. (2016). "Group Emotion and Group Understanding." In M. Brady and M. Fricker (eds.), *The Epistemic Life of Groups: Essays in the Epistemology of Collectives*, Oxford: Oxford University Press, pp. 95-110.

Bullock, E. (2018). "Knowing and Not-Knowing For Your Own Good: The Limits of Epistemic Paternalism." *Journal of Applied Philosophy* 35(2): 433-447.

Carter, J. A. (2015). "Group Knowledge and Epistemic Defeat." *Ergo* 28(2): 711-735.

Carter, J. A. and Czarnecki, B. (2016). "Extended Knowledge-How." *Erkenntnis* 81(2): 259-273.

Carter, J. A., and Gordon, E. (2014). "Objectual Understanding and the Value Problem." *American Philosophical Quarterly* 51: 1-13.

Carter, J. A. and Pritchard, D. (2015a). "Knowing How and Cognitive Achievement." *Philosophy and Phenomenological Research* 91(1): 181-199.

Carter, J. A. and Pritchard, D. (2015b). "Knowledge-how and Epistemic Value." *Australasian Journal of Philosophy* 93(4): 799-816.

Cassam, Q. (2016). "Vice Epistemology." *The Monist* 99: 159-180.

Chi, M. T. H. (2006). "Two Approaches to the Study of Experts' Characteristics." In K. Anders Ericsson (ed.), *The Cambridge Handbook of Expertise and Expert Performance*, New York: Cambridge University Press, pp. 21-30.

Coady, D. (2006). "When Experts Disagree." *Episteme* 3(1-2): 68-79.

Coady, D. (2012). *What to Believe Now: Applying Epistemology to Contemporary Issues*. Malden (MA): Wiley-Blackwell.

Code, L. (1991). *What Can She Know?* Ithaca (NY): Cornell University Press.

Collins, H. and Evans, R. (2007). *Rethinking Expertise*. Chicago: University of Chicago Press.

Collins, P. (1990). *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. New York: Routledge.

Constantin, J. and Grundmann, T. (2018). "Epistemic Authority: Preemption through Source-sensitive Defeat." *Synthese*. doi: 10.1007/s11229-018-01923-x.

Cordell, S. (2017). "Group Virtues: No Great Leap Forward with Collectivism." *Res Publica* 23(1): 43-59.

Croce, M. (2018a). "Expert-oriented Abilities vs. Novice-oriented Abilities: An Alternative Account of Epistemic Authority." *Episteme* 15(4): 476-498.

Croce, M. (2018b). "Epistemic Paternalism and the Service Conception of Epistemic Authority." *Metaphilosophy* 49(3): 305-327.

Croce, M. (2018c). "Il potenziale educativo degli esemplari intellettuali." *Ethics and Politics* 20(2): 143-162.

Croce, M. (2019a). "On What It Takes To Be An Expert." *The Philosophical Quarterly* 69: 1-21.

Croce, M. (2019b). "Exemplarism in Moral Education: Problems with Applicability and Indoctrination." *Journal of Moral Education* 48(3): 291-302.

Croce, M. (2020a). "Moral Exemplars in Education: A Liberal Account." *Ethics and Education* 15(2): 186-199.

Croce, M. (2020b). "Epistemic Paternalism, Personal Sovereignty, and One's Own Good." In G. Axtell and A. Bernal (eds.), *Epistemic Paternalism: Conceptions, Justifications, and Implications*. Lanham: Rowman & Littlefield.

Croce, M. (2020c). "The Right to Know and the Duty to Inform: A Lesson from the Italian Experience with Covid-19." Post for *Open for Debate Blog*. Available at: https://blogs.cardiff.ac.uk/openfordebate/2020/03/23/the-right-to-know-and-the-duty-to-inform-a-lesson-from-the-italian-experience-with-covid-19/.

Croce, M., and Piazza, T. Forthcoming. "Misinformation and Intentional Deception: A Novel Account of Fake News." In N. Snow and M. S. Vaccarezza (eds.), *Virtues, Democracy, and Online Media: Ethical and Epistemic Issues*, London: Routledge.

Croce, M., and Vaccarezza, M. S. (2017). "Educating through Exemplars: Alternative Paths to Virtue." *Theory and Research in Education* 15(1): 5-19.

Delarivière, S. (2020). "Collective Understanding—A Conceptual Defense for When Groups Should Be Regarded as Epistemic Agents with Understanding." *Avant* 11(2): 1-20.

Dormandy, K. (2018). "Epistemic Authority: Preemption or Proper Basing?" *Erkenntnis* 83(4): 773-791.

Dreyfus, H. and Dreyfus, S. (1991). "Towards A Phenomenology of Ethical Expertise." *Human Studies* 14: 229-250.

Driver, J. (2013). "Moral Expertise, Judgment, Practice, Analysis." *Social Philosophy and Policy* 30: 280-296.

Dworkin, R. (2017). "Paternalism." In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/paternalism/.

Elga, A. (2007). "Reflection and Disagreement." *Nous* 41(3): 478-502.

Elgin, C. (2007). "Understanding and the Facts." *Philosophical Studies* 132: 33-42.

Elgin, C. (2017). *True Enough*. Cambridge (MA): The MIT Press.

Feinberg, J. (1989). *The Moral Limits of the Criminal Law, Volume 3: Harm to Self.* Oxford: Oxford University Press.

Fricker, E. (2006). "Testimony and Epistemic Autonomy." In J. Lackey and E. Sosa (eds.), *The Epistemology of Testimony*, Oxford: Oxford University Press, pp. 225-250.

Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.

Fricker, M. (2010). "Can There Be Institutional Virtues?" In T. S. Gendler and J. Hawthorne (eds.), *Oxford Readings in Epistemology, Vol. 3*, Oxford: Oxford University Press, pp. 235-252.

Fricker, M. (2013). "Epistemic Justice as a Condition of Political Freedom?" *Synthese* 190: 1317-1332.

Fuller, S. (1988). *Social Epistemology*. Bloomington: Indiana University Press.

Fuller, S. (2018). *Post-Truth: Knowledge as a Power Game.* London: Anthem Press.

Gasparov, I. (2017). "Epistemic Authority and Autonomy of the Epistemic Subject." *Epistemology and Philosophy of Science* 53(3): 108-122.

Gelfert, A. (2018). "Fake News: A Definition." *Informal Logic* 38: 84-117.

Gilbert, M. (2013). *Joint Commitment: How We Make the Social World*. Oxford: Oxford University Press.

Goldberg, S. (2010). *Relying on Others: An Essay in Epistemology*. Oxford: Oxford University Press.

Goldberg, S. (2019). "Against Epistemic Partiality in Friendship: Value-Reflecting Reasons." *Philosophical Studies* 176(8): 2221-2242.

Goldman, A. (1978). "Epistemics: The Regulative Theory of Cognition." *Journal of Philosophy* 75: 509-523.

Goldman, A. (1991). "Epistemic Paternalism: Communication Control in Law and Society." *The Journal of Philosophy* 88(3): 113-131.

Goldman, A. (1999). *Knowledge in a Social World.* Oxford: Oxford University Press.

Goldman, A. (2001). "Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63: 85-110.

Goldman, A. (2018). "Expertise." *Topoi* 37(1): 3-10.

Goldman, A. and Blanchard, T. (2015). "Social Epistemology." In E. N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/sum2018/epistemology-social/.

Gordon, E. (2017). "Social Epistemology and the Acquisition of Understanding." In S. Grimm, C. Baumberger, and S. Ammonn (eds.), *Explaining Understanding: New*

Perspectives from Epistemology and the Philosophy of Science, New York: Routledge, pp. 293-317.

Grasswick, H. (2018). "Feminist Social Epistemology." In E. N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/fall2018/entries/epistemology-social/.

Greco, J. (2015). "Testimonial Knowledge and the Flow of Information." In D. K. Henderson and J. Greco (eds.), *Epistemic Evaluation: Purposeful Epistemology*, Oxford: Oxford University Press, pp. 274-290.

Grimm, S. (2014). "Understanding as Knowledge of Causes." In A. Fairweather (ed.), *Virtue Epistemology Naturalized: Bridges Between Virtue Epistemology and Philosophy of Science*, New York: Springer, pp. 329-346.

Grimm, S. (2018). "Understanding as an Intellectual Virtue." In H. Battaly (ed.), *The Routledge Handbook of Virtue Epistemology*, New York: Routledge, pp. 340-351.

Grundmann, T. Forthcoming. "Preemptive Authority: The Challenge from Outrageous Expert Judgment." *Episteme*.

Haack, S. (2019). "Post Post-Truth: Are We There Yet?" *Theoria* 85(4): 258-275.

Han, H., Kim, J., Jeong, C. and Cohen, G. (2017). "Attainable and Relevant Moral Exemplars are More Effective than Extraordinary Exemplars in Promoting Voluntary Service Engagement." *Frontiers in Psychology* 8: 283.

Harding, S. (1991). *Whose Science? Whose Knowledge? Thinking from Women's Lives*. Ithaca: Cornell University Press.

Hardwig, J. (1985). "Epistemic Dependence." *The Journal of Philosophy* 82: 335-349.

Hart, H. L. A. (1990). "Commands and Authoritative Legal Reasons." In J. Raz (ed.), *Authority*, New York: New York University Press, pp. 92-114.

Hawley, K. (2012). *Trust: A Very Short Introduction*. Oxford: Oxford University Press.

Hawley, K. (2014). "Partiality and Prejudice in Trusting." *Synthese* 191(9): 2029-2045.

Hills, A. (2009). "Moral Testimony and Moral Epistemology." *Ethics* 120: 94-127.

Hills, A. (2016). "Understanding Why." *Nous* 50(4): 661-688.

Jäger, C. (2016). "Epistemic Authority, Preemptive Reasons, and Understanding." *Episteme* 13(2): 167-185.

Kallestrup, J. (2016). "Group Virtue Epistemology." *Synthese*. doi: 10.1007/s11229-016-1225-7.

Kawall J. (2013). "Friendship and Epistemic Norms." *Philosophical Studies* 165(2): 349-370.

Kelp, C. (2015). "Understanding Phenomena." *Synthese* 192(12): 3799-3816.

Keren, A. (2014a). "Trust and Belief: A Preemptive Reasons Account." *Synthese* 191: 2593-2615.

Keren, A. (2014b). "Zagzebski on Authority and Pre-emption in the Domain of Belief." *European Journal for Philosophy of Religion* 6: 61-76.

Kitcher, P. (1990). "The Division of Cognitive Labor." *The Journal of Philosophy* 87: 5-22.

Korsgaard, M. T. (2019). "Exploring the Role of Exemplarity in Education: Two Dimensions of The Teacher's Task." *Ethics and Education* 14(3): 271-284.

Kotzee, B., Carter, J. A., and Siegel, H. (2019). "Educating for Intellectual Virtue: A Critique from Action Guidance." *Episteme*. doi: 10.1017/epi.2019.10

Kristjánsson, K. (2015). *Aristotelian Character Education.* London: Routledge.

Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding.* New York: Cambridge University Press.

Kvanvig, J. (2009) "The Value of Understanding." In A. Haddock, A. Millar, and D. Pritchard (eds.), *Epistemic Value*, New York: Cambridge University Press, pp. 95-111.

Lackey, J. (2011). "Assertion and Isolated Second-Hand Knowledge." In J. Brown and H. Cappelen (eds.), *Assertion: New Philosophical Essays*, Oxford: Oxford University Press, pp. 251-275.

Lackey, J. (2012). "Group Knowledge Attributions." In J. Brown amd M. Gerken (eds.), *New Essays on Knowledge Ascriptions*, Oxford: Oxford University Press, pp. 243-269.

Lackey, J. (2014). "Socially Extended Knowledge." *Philosophical Issues* 24(1): 282-298.

Lackey, J. (2016). "Assertion and Expertise." *Philosophy and Phenomenological Research* 92(2): 509-517.

Lackey, J. (2018). "Experts and Peer Disagreement." In M. Benton, J. Hawthorne, and D. Rabinowitz (eds.), *Knowledge, Belief, and God: New Insights in Religious Epistemology*, Oxford: Oxford University Press, pp. 228-245.

Lackey, J. Forthcoming. *The Epistemology of Groups.* New York: Oxford University Press.

Lahroodi, R. (2007). "Collective Epistemic Virtues." *Social Epistemology* 21(3): 281-297.

Lasoneen-Aarnio, M. (2020). "Enkrasia or Evidentialism? Learning to Love Mismatch." *Philosophical Studies* 177: 597-632.

Lefevere, M. and Schliesser, E. (2014). "Private Epistemic Virtue, Public Vices: Moral Responsibility in the Policy Sciences." In C. Martini and M. Bouman (eds.), *Experts and Consensus in Social Science*, Dordrecht: Springer, pp. 275-295.

Longino, H. (1990). *Science as Social Knowledge*. Princeton (NJ): Princeton University

Press.

Lynch, M. P. (2016). *The Internet of Us: Knowing More and Understanding Less in the Age of Big Data*. New York: Norton & Company.

Majdik, Z. P. and Keith, W. M. (2011). "The Problem of Pluralistic Expertise: A Wittgensteinian Approach to the Rhetorical Basis of Expertise." *Social Epistemology* 25(3): 275-290.

Malfatti, F. I. (2019a). "On Understanding and Testimony." *Erkenntnis*. doi: 10.1007/s10670-019-00157-8.

Malfatti, F. I. (2019b). "Can Testimony Generate Understanding?" *Social Epistemology* 33(6): 477-490.

Masui, K. and Hashimoto, J. (1933). *Sexing Baby Chicks*. Vancouver: Journal Printing Company.

McGlynn, A. (2019). "Epistemic Objectification as the Primary Harm of Testimonial Injustice." *Episteme*. doi: 10.1017/epi.2019.9.

Medina, J. (2013). *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imagination*. Oxford. Oxford University Press.

Meeker, K. (2013). *Hume's Radical Scepticism and the Fate of Naturalized Epistemology*. New York: Palgrave MacMillan.

Mlodinow, L. (2008). *The Drunkard's Walk: How Randomness Rules Our Lives*. New York: Random House.

Muldoon, R. (2017). "Diversity, Rationality, and the Division of Cognitive Labor." In T. Boyer-Kassem, C. Mayo-Wilson, and M. Weisberg (eds.), *Scientific Collaboration and Collective Knowledge: New Essays*, Oxford: Oxford University Press, pp. 78-91.

Narvaez, D. and Lapsey, D. K. (2005). "The Psychological Foundations of Everyday Morality and Moral Expertise." In D. Lapsley and C. Power (eds.), *Character Psychology and Character Education*, Notre Dame (IN): University of Notre Dame Press, pp. 140-165.

Nichols, T. (2017). *The Death of Expertise: The Campaign against Established Knowledge and Why It Matters*. New York: Oxford University Press.

Nguyen, T. (2018). "Echo Chambers and Epistemic Bubbles." *Episteme.* doi: 10.1017/epi.2018.32.

Penco, C. (2014). "Between Knowing How and Knowing That." In J. Dutant, D. Fassio, and A. Meylan (eds.), *Liber Amicorum Pascal Engel*. Université de Genève. http://www.unige.ch/lettres/philo/publications/engel/liberamicorum/.

Porter, S. (2016). "A Therapeutic Approach to Intellectual Virtue Formation in the Classroom." In J. Baehr (ed.), *Intellectual Virtues and Education: Essays in Applied Virtue Epistemology,* London: Routledge, pp. 221-239.

Pritchard, D. (2009). "Knowledge, Understanding and Epistemic Value." *Royal Institute of Philosophy Supplement* 64: 19-43.

Pritchard, D. (2010). "Knowledge and Understanding." In D. Pritchard, A. Millar, and A. Haddock, *The Nature and Value of Knowledge*, Oxford: Oxford University Press, pp. 3-88.

Pritchard, D. (2013a). "Epistemic Virtue and the Epistemology of Education." *Journal of Philosophy of Education* 47: 236-247.

Pritchard, D. (2013b). "Epistemic Paternalism and Epistemic Value." *Philosophical Inquiries* 1(2): 9-37.

Pritchard, D. (2015). "Epistemic Dependence." *Philosophical Perspectives* 30: 305-324.

Pritchard, D. (2018). "Neuromedia and the Epistemology of Education." *Metaphilosophy* 49(3): 328-349.

Pritchard, D. (2019). "Intellectual Virtues and the Epistemic Value of Truth." *Synthese*. doi: 10.1007/s11229-019-02418-z.

Pritchard, D. Forthcoming. "Good News, Bad News, Fake News." In S. Bernecker, A. Floweree, and T. Grundmann (eds.), *The Epistemology of Fake News*, Oxford: Oxford University Press.

Quast, C. (2018). "Expertise: A Practical Explanation." *Topoi* 37(1): 11-27.

Raz, J. (1979). *The Authority of Law: Essays on Law and Morality*. Oxford: Clarendon Press.

Raz, J. (1986) *The Morality of Freedom*. Oxford: Clarendon Press.

Raz, J. (2006). "The Problem of Authority: Revisiting the Service Conception." *Minnesota Law Review* 90: 1003-44.

Roberts, R. and Wood, W. J. (2007). *Intellectual Virtues: An Essay in Regulative Epistemology*. Oxford: Clarendon Press.

Rorty, A. (1996). "Descartes and Spinoza on Epistemological Egalitarianism." *History of Philosophy Quarterly* 13(1): 35-53.

Ryan, S. (2011). "Wisdom, Knowledge, and Rationality." *Acta Analitica* 27: 99-112.

Ryan, S. (2016). "Paternalism: An Analysis." *Utilitas* 28(2): 123-135.

Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson.

Scholz, O. R. (2009). "Experts: What They Are and How We Recognize Them—A Discussion of Alvin Goldman's Views." *Grazer Philosophische Studien* 79: 187-205.

Scholz, O. R. (2018). “Symptoms of Expertise: Knowledge, Understanding and Other Cognitive Goods.” *Topoi* 37(1): 29-37.

Sherman, N. (2016). “There’s No (Testimonial) Justice: Why Pursuit of a Virtue is Not the Solution to Epistemic Justice.” *Social Epistemology* 30(3): 229-250.

Siegel, H. (2018). *Education’s Epistemology: Rationality, Diversity, and Critical Thinking*. New York: Oxford University Press.

Stanley, J. 2011. “Knowing (How).” *Nous* 45(2): 207-238.

Stanley, J. and Williamson, T. (2001). “Knowing How.” *The Journal of Philosophy* 98(8): 411-444.

Steutel, J. and Carr, D. (1999). “Virtue Ethics and the Virtue Approach to Moral Education.” In D. Carr and J. Steutel (eds.), *Virtue Ethics and Moral Education*, London: Routledge, pp. 3-18.

Stichter, M. (2015). “Philosophical and Psychological Accounts of Expertise and Experts.” *Humana.mente* 28: 105-128.

Stichter, M. (2018). *The Skillfulness of Virtue: Improving Our Moral and Epistemic Lives*. New York: Cambridge University Press.

Strevens, M. (2003). “The Role of the Priority Rule in Science.” *Journal of Philosophy* 100: 55-79.

Stroud, S. (2006). “Epistemic Partiality in Friendship.” *Ethics* 116(3): 498-524.

Sullivan, E. and Alfano, M. (2019). “Negative Epistemic Exemplars.” In N. Sherman and S. Goguen (eds.), *Overcoming Epistemic Injustice: Social and Psychological Perspectives*, Lanham (MD): Rowman & Littlefield.

Sunstein, C. (2017). *#Republic: Divided Democracy in the Age of Social Media*. Princeton (NJ): Princeton University Press.

Szutta, N. (2019). “Exemplarist Moral Theory—Some Pros and Cons.” *Journal of Moral Education* 48(3): 280-290.

Tanesini, A. (2016). “Teaching Virtues: Changing Attitudes.” *Logos and Episteme* 7(4): 503-527.

Tuomela, R. (1992). “Group Beliefs.” *Synthese* 91: 285-318.

Vaccarezza, M. S. and Niccoli, A. (2019). “The Dark Side of The Exceptional: On Moral Exemplars, Character Education, and Negative Emotions.” *Journal of Moral Education* 48(3): 332-345.

Ward, T. (2017). “Expert Testimony, Law, and Epistemic Authority.” *Journal of Applied Philosophy* 34(2): 263-277.

Watson, J. C. (2018). “The Shoulders of Giants: A Case for Non-veritism about Expert Authority.” *Topoi* 37(1): 39-53.

Watson, L. (2015). “What is Inquisitiveness.” *American Philosophical Quarterly* 52(3): 273-288.

Watson, L. (2016). “Why Should We Educate for Inquisitiveness.” In J. Baehr (ed.), *Intellectual Virtues and Education: Essays in Applied Virtue Epistemology*, London: Routledge, pp. 38-53.

Watson, L. (2019). “Educating for Inquisitiveness: A Case Against Exemplarism for Intellectual Character Education.” *Journal of Moral Education* 48(3): 303-315.

Whitcomb, D. (2010). “Wisdom.” In S. Bernecker and D. Pritchard (eds.), *Routledge Companion to Epistemology*, London: Routledge, pp. 95-105.

Wilkenfield, D. (2013). “Understanding as Representation Manipulability.” *Synthese* 190(6): 997-1016.

Wilkenfield, D. A., Plunkett, D., and Lombrozo, T. (2016). “Depth and Deference: When and Why We Attribute Understanding.” *Philosophical Studies* 173(2): 373-393.

Williams, B. (2002). *Truth and Truthfulness: An Essay in Genealogy*. Princeton (NJ): Princeton University Press.

Wolf, S. (1982). “Moral Saints.” *The Journal of Philosophy* 79(8): 419-439.

Zagzebski, L. (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge.* Cambridge: Cambridge University Press.

Zagzebski, L. (2003). “Emotion and Moral Judgment.” *Philosophy and Phenomenological Research* 66(1): 104-124.

Zagzebski, L. (2008). *On Epistemology*. Belmont: Wadsworth.

Zagzebski, L. (2012). *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief*. Oxford: Oxford University Press.

Zagzebski, L. (2014). “Epistemic Authority and Its Critics.” *European Journal for Philosophy of Religion* 6: 169-187.

Zagzebski, L. (2016). “Replies to Christoph Jäger and Elizabeth Fricker.” *Episteme* 13(2): 187-194.

Zagzebski, L. (2017). *Exemplarist Moral Theory*. New York: Oxford University Press.

Zollman, K. J. S. (2010). “The Epistemic Benefit of Transient Diversity.” *Erkenntnis* 72: 17-35.